Efficient Craig Interpolation for Linear Diophantine (Dis)Equations and Linear Modular Equations

Himanshu Jain[†] **Edmund M. Clarke**[†] **Orna Grumberg**[*]

February 2008
CMU-CS-08-102

School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213

[†] School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, USA
[*] Department of Computer Science, Technion - Israel Institute of Technology

This research was sponsored by the Gigascale Systems Research Center (GSRC), Semiconductor Research Corporation (SRC), the National Science Foundation (NSF), the Office of Naval Research (ONR), the Naval Research Laboratory (NRL), the Defense Advanced Research Projects Agency (DARPA), the Army Research Office (ARO), and the General Motors Collaborative Research Lab at CMU. The views and conclusions contained in this document are those of the author and should not be interpreted as representing the official policies, either expressed or implied, of GSRC, SRC, NSF, ONR, NRL, DARPA, ARO, GM, or the U.S. government.

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
FEB 2008		**00-00-2008 to 00-00-2008**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Efficient Craig Interpolation for Linear Diophantine (Dis)Equations and Linear Modular Equations	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Carnegie Mellon University,School of Computer Science,Pittsburgh,PA,15213**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The use of Craig interpolants has enabled the development of powerful hardware and software model checking techniques. Efficient algorithms are known for computing interpolants in rational and real linear arithmetic. We focus on subsets of integer linear arithmetic. Our main results are polynomial time algorithms for obtaining proofs of unsatisfiability and interpolants for conjunctions of linear diophantine equations linear modular equations (linear congruences), and linear diophantine disequations. We show the utility of the proposed interpolation algorithms for discovering modular/divisibility predicates in a counterexample guided abstraction refinement (CEGAR) framework. This has enabled verification of simple programs that cannot be checked using existing CEGAR based model checkers.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**39**	

Keywords: Craig Interpolation, Proofs of Unsatisfiability, Linear Diophantine Equations, Linear Modular Equations (Linear Congruences), Linear Diophantine Disequations, Abstraction Refinement

Abstract

The use of Craig interpolants has enabled the development of powerful hardware and software model checking techniques. Efficient algorithms are known for computing interpolants in rational and real linear arithmetic. We focus on subsets of integer linear arithmetic. Our main results are polynomial time algorithms for obtaining proofs of unsatisfiability and interpolants for conjunctions of linear diophantine equations, linear modular equations (linear congruences), and linear diophantine disequations. We show the utility of the proposed interpolation algorithms for discovering *modular/divisibility* predicates in a counterexample guided abstraction refinement (CEGAR) framework. This has enabled verification of simple programs that cannot be checked using existing CEGAR based model checkers.

1 Introduction

The use of Craig interpolation [12] has led to powerful hardware [23] and software [17] model checking techniques. In [23] the idea of interpolation is used for obtaining over-approximations of the reachable set of states without using the costly image computation (existential quantification) operations. In [17, 18] interpolants are used for finding the *right* set of predicates in order to rule out *spurious counterexamples*. An interpolating theorem prover performs the task of finding the interpolants. Such provers are available for various theories such as propositional logic, rational and real linear arithmetic and equality with uninterpreted functions [24, 33, 19, 18, 28, 20, 10].

Efficient algorithms are known for computing interpolants in rational and real linear arithmetic [24, 28, 10]. Linear arithmetic formulas where all variables are constrained to be integers are said to be formulas in *(pure) integer linear arithmetic* or $LA(\mathbb{Z})$, where $\mathbb{Z}$ is the set of integers. There are no known efficient algorithms for computing interpolants for formulas in $LA(\mathbb{Z})$. This is expected because checking the satisfiability of conjunctions of atomic formulas in $LA(\mathbb{Z})$ is itself NP-hard. We show that for various *subsets* of $LA(\mathbb{Z})$ one can compute proofs of unsatisfiability and interpolants efficiently.

Informally, a linear equation where all variables are integer variables is said to be a *linear diophantine equation (LDE)*. A *linear modular equation (LME)* or a *linear congruence* over integer variables is a type of linear equation that expresses divisibility relationships. A *system* of LDEs (LMEs) denotes conjunctions of LDEs (LMEs). Both LDEs and LMEs arise naturally in program verification when modeling assignments and conditional statements as logical formulas. These subsets of $LA(\mathbb{Z})$ are also known to be tractable, that is, polynomial time algorithms are known for deciding systems of LDEs and LMEs. We study the interpolation problem for LDEs and LMEs.

Given formulas F, G such that $F \wedge G$ is unsatisfiable. An interpolant for the pair (F, G) is a formula $I(F, G)$ with the following properties: (i) F implies $I(F, G)$, (ii) $I(F, G) \wedge G$ is unsatisfiable, and (iii) $I(F, G)$ refers only to the common variables of F and G. This paper presents the following new results.

- F, G denote a system of LDEs: We show that $I(F, G)$ can be obtained in polynomial time by using a proof of unsatisfiability of $F \wedge G$. The interpolant can be either a LDE or a LME. This is because in some cases there is no $I(F, G)$ that is a LDE. In these cases, however, there is always an $I(F, G)$ in the form of a LME. (Section 3)

- F, G denote a system of LMEs: We obtain $I(F, G)$ in polynomial time by using a proof of unsatisfiability of $F \wedge G$. We can ensure that $I(F, G)$ is a LME. (Section 4)

- Let S denote an unsatisfiable system of LDEs. The proof of unsatisfiability of S can be obtained in polynomial time by using the *Hermite Normal Form* of S (represented in matrix form). A system of LMEs R can be reduced to an equi-satisfiable system of LDEs R'. The proof of unsatisfiability for R is easily obtained from the proof of unsatisfiability of R'. (Section 5)

- Let S denote a system of LDEs. We show that if S has an integral solution, then every LDE that is implied by S, can be obtained by a linear combination of equations in S. We show that S is *convex* [25], that is, if S implies a disjunction of LDEs, then it implies one of the equations in the disjunction. In contrast, conjunctions of atomic formulas in $LA(\mathbb{Z})$ are not convex due to inequalities [25]. These results help in efficiently dealing with *linear diophantine disequations (LDDs)*. (Section 6)

- Let $S = S_1 \wedge S_2$, where S_1 is a system of LDEs, while S_2 is a system of LDDs. We say that S is a system of LDEs+LDDs. We show that S has no integral solution if and only if $S_1 \wedge S_2$ has no rational

solution or S_1 has no integral solution. This gives a polynomial time decision procedure for checking if S has an integral solution. If S has no integral solution, then the proof of unsatisfiability of S can be obtained in polynomial time. (Section 6)

- F, G denote a system of LDEs+LDDs: We show $I(F, G)$ can be obtained in polynomial time. The interpolant can be an LDE, an LDD, or an LME. (Section 6)

- We show the utility of our interpolation algorithms in counterexample guided abstraction refinement (CEGAR) based verification [11]. Our interpolation algorithm is effective at discovering *modular/divisibility predicates*, such as $3x + y + 2z \equiv 1 \ (mod\ 4)$, from spurious counterexamples. This has allowed us to verify programs that cannot be verified by existing hardware and software model checkers. (Section 7)

Polynomial time algorithms are known for solving (deciding) a system of LDEs [29, 7] and LMEs (by reduction to LDEs) over integers. We do not give any new algorithms for solving a system of LDEs or LMEs. Instead we focus on obtaining proofs of unsatisfiability and interpolants for systems of LDEs, LMEs, LDEs+LDDs. We only consider conjunctions of LDEs, LMEs, LDEs+LDDs. Interpolants for any (unsatisfiable) Boolean combinations of LDEs can also be obtained by calling the interpolation algorithm for conjunctions of LDEs+LDDs multiple times in a satisfiability modulo theory (SMT) framework [10]. However, computing interpolants for Boolean combinations of LMEs is difficult. This is due to linear modular disequations (LMDs). We can show that even the decision problem for conjunctions of LMDs is NP-hard.

All proofs are present in the appendix of this paper.

1.1 Related work

It is known that Presburger arithmetic (PA) allows quantifier elimination [26]. Kapur et al. [19] show that a recursively enumerable theory allows quantifier-free interpolants if and only if it allows quantifier elimination. The systems of LDEs, LMEs, LDEs+LDDs are subsets of PA. Thus, the existence of quantifier-free interpolants for these systems follows from [19]. However, quantifier elimination for PA has an exponential complexity and does not immediately yield efficient algorithms for computing interpolants. We give polynomial time algorithms for computing proofs of unsatisfiability and interpolants for systems (conjunctions) of LDEs, LMEs, LDEs+LDDs.

Let S_1, S_2 denote conjunctions of atomic formulas in $LA(\mathbb{Z})$. Suppose $S_1 \wedge S_2$ is unsatisfiable. Pudlak [27] shows how to compute an interpolant for (S_1, S_2) by using a *cutting-plane* (CP) proof of unsatisfiability. The CP proof system is a sound and complete way of proving unsatisfiability of conjunctions of atomic formulas in $LA(\mathbb{Z})$. However, a CP proof for a formula can be exponential in the size of the formula. Pudlak does not provide any guarantee on the size of CP proofs for a system of LDEs or LMEs. Our results show that polynomially sized proofs of unsatisfiability and interpolants can be obtained for systems of LDEs, LMEs and LDEs+LDDs.

McMillan [24] shows how to compute interpolants in the combined theory of rational linear arithmetic $LA(\mathbb{Q})$ and equality with uninterpreted functions $\mathcal{EUF}$ by using proofs of unsatisfiability. Rybalchenko and Sofronie-Stokkermans [28] show how to compute interpolants in combined $LA(\mathbb{Q})$, $\mathcal{EUF}$ and real linear arithmetic $LA(\mathbb{R})$ by using linear programming solvers in a black-box fashion. The key idea in [28] is to use an extension of Farkas lemma [29] to reduce the interpolation problem to constraint solving in $LA(\mathbb{Q})$ and $LA(\mathbb{R})$. Cimatti et al. [10] show how to compute interpolants in a satisfiability modulo theory (SMT) framework for $LA(\mathbb{Q})$, rational difference logic fragment and $\mathcal{EUF}$. By making use of state-of-the-art SMT

algorithms [14] they obtain significant improvements over existing interpolation tools for $LA(\mathbb{Q})$ and $\mathcal{EUF}$. Yorsh and Musuvathi [33] give a Nelson-Oppen [25] style method for generating interpolants in a combined theory by using the interpolation procedures for individual theories. Kroening and Weissenbacher [20] show how a bit-level proof can be lifted to a word-level proof of unsatisfiability (and interpolants) for equality logic.

To the best of our knowledge the work in [24, 33, 28, 20, 10] is not complete for computing interpolants in $LA(\mathbb{Z})$ or its subsets such as LDEs, LMEs, LDEs+LDDs. That is, the work in [24, 33, 28, 20, 10] cannot compute interpolants for formulas that are satisfiable over rationals but unsatisfiable over integers. Such formulas can arise in both hardware and software verification. We give sound and complete polynomial time algorithms for computing interpolants for conjunctions of LDEs, LMEs, LDEs+LDDs. Efficient interpolation algorithms for LDEs, LMEs, LDEs+LDDs are also crucial in order to develop practical interpolating theorem provers for $LA(\mathbb{Z})$ and bit-vector arithmetic [13, 6, 5, 15, 21, 9, 16, 8].

2 Notation and preliminaries

We use capital letters $A, B, C, X, Y, Z, \ldots$ to denote matrices and formulas. A matrix M is *integral (rational)* iff all elements of M are integers (rationals). For a matrix M with m rows and n columns we say that the size of M is $m \times n$. A *row vector* is a matrix with a single row. A *column vector* is a matrix with a single column. We sometimes identify a matrix M of size 1×1 by its only element. If A, B are matrices, then AB denotes matrix multiplication. We assume that all matrix operations are well defined in this paper. For example, when we write AB without specifying the sizes of matrices A, B, it is assumed that the number of columns in A equals the number of rows in B.

For any rational numbers α and β, $\alpha|\beta$ if and only if, α divides β, that is, if and only if $\beta = \lambda\alpha$ for some integer λ. We say that α is equivalent to β *modulo* γ written as $\alpha \equiv \beta \ (mod\ \gamma)$ if and only if $\gamma|(\alpha - \beta)$. We say γ is the *modulus* of the equation $\alpha \equiv \beta \ (mod\ \gamma)$. We allow α, β, γ to be rational numbers. If $\alpha_1, \ldots, \alpha_n$ are rational numbers, not all equal to 0, then the largest rational number γ dividing each of $\alpha_1, \ldots, \alpha_n$ exists [29], and is called the *greatest common divisor*, or *gcd* of $\alpha_1, \ldots, \alpha_n$ denoted by $gcd(\alpha_1, \ldots, \alpha_n)$. We assume that gcd is always positive.

Basic Properties of Modular Arithmetic: Let a, b, c, d, m be rational numbers.
P1. $a \equiv a \ (mod\ m)$ (reflexivity).
P2. $a \equiv b \ (mod\ m)$ implies $b \equiv a \ (mod\ m)$ (symmetry).
P3. $a \equiv b \ (mod\ m)$ and $b \equiv c \ (mod\ m)$ imply $a \equiv c \ (mod\ m)$ (transitivity).
P4. If $a \equiv b \ (mod\ m)$, $c \equiv d \ (mod\ m)$, and x, y are integers, then $ax + cy \equiv bx + dy \ (mod\ m)$ (integer linear combination).
P5. If $c > 0$ then $a \equiv b \ (mod\ m)$ if, and only if, $ac \equiv bc \ (mod\ mc)$.
P6. If $a = b$, then $a \equiv b \ (mod\ m)$ for any m.

Example 1 Observe that $x \equiv 0 \ (mod\ 1)$ for any integer x. Also observe from P5 (with $c = 2$) that $\frac{1}{2}x \equiv 0 \ (mod\ 1)$ if and only if $x \equiv 0 \ (mod\ 2)$.

A *linear diophantine equation (LDE)* is a linear equation $c_1x_1 + \ldots + c_nx_n = c_0$, where $x_1, \ldots, x_n$ are integer variables and $c_0, \ldots, c_n$ are rational numbers. A variable x_i is said to *occur* in the LDE if $c_i \neq 0$. We denote a system of m LDEs in a matrix form as $CX = D$, where C denotes an $m \times n$ matrix of rationals, X denotes a column vector of n integer variables and D denotes a column vector of m rationals. When we write a (single) LDE in the form $CX = D$, it is implicitly assumed that the sizes of C, X, D are of the form

$1 \times n, n \times 1, 1 \times 1$, respectively. A variable is said to *occur* in a system of LDEs if it occurs in at least one of the LDEs in the given system of LDEs.

A *linear modular equation (LME)* has the form $c_1x_1 + \ldots + c_nx_n \equiv c_0 \ (mod\ l)$, where $x_1, \ldots, x_n$ are integer variables, $c_0, \ldots, c_n$ are rational numbers, and l is a rational number. We call l the modulus of the LME. Allowing l to be a rational number allows for simpler proofs and covers the case when l is an integer. For brevity, we write a LME $t \equiv c \ (mod\ l)$ by $t \equiv_l c$. A variable x_i is said to *occur* in a LME if l does not divide c_i.

A *system* of LDEs (LMEs) denotes conjunctions of LDEs(LMEs). If F, G are a system of LDEs (LMEs), then $F \wedge G$ is also a system of LDEs (LMEs).

2.1 Craig Interpolants

Given two logical formulas F and G in a theory $\mathcal{T}$ such that $F \wedge G$ is unsatisfiable in $\mathcal{T}$. An *interpolant* I for the ordered pair (F, G) is a formula such that
(1) $F \Rightarrow I$ in $\mathcal{T}$
(2) $I \wedge G$ is unsatisfiable in $\mathcal{T}$
(3) I refers to only the common variables of A and B.
The interpolant I can contain symbols that are interpreted by $\mathcal{T}$. In this paper such symbols will be one of the following: addition ($+$), equality ($=$), modular equality for some rational number m ($\equiv_m$), disequality ($\neq$), and multiplication by a rational number ($\times$). The exact set of interpreted symbols in the interpolant depends on $\mathcal{T}$.

3 System of linear diophantine equations (LDEs)

In this section we discuss proofs of unsatisfiability and interpolation algorithm for LDEs. The following theorem from [29] gives a necessary and sufficient condition for a system of LDEs to have an integral solution.

Theorem 1 *(Schrijver [29]) A system of LDEs $CX = D$ has no integral solution for X, if and only if there exists a rational row vector R such that RC is integral and RD is not an integer.*

Definition 1 *We say a system of LDEs $CX = D$ is* **unsatisfiable** *if it has no integral solution for X. For a system of LDEs $CX = D$ a* **proof of unsatisfiability** *is a rational row vector R such that RC is integral and RD is not an integer.*

In section 5 we describe how a proof of unsatisfiability R can be obtained in polynomial time for an unsatisfiable system of LDEs. (We show in the appendix I that R can be converted to a polynomially sized proof in a *cutting-plane* proof system [29, 7].)

Example 2 Consider the system of LDEs $CX = D$ and a proof of unsatisfiability R:

$$CX = D := \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 2 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix} \qquad \begin{array}{lcl} R & = & [\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}] \\ RC & = & [0, 2, 1] \\ RD & = & \frac{3}{2} \end{array}$$

Example 3 Consider the system of LDEs $CX = D$ and a proof of unsatisfiability R:

$$CX = D := \begin{bmatrix} 1 & -2 & 0 \\ 1 & 0 & -2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad \begin{array}{lcl} R & = & [\frac{1}{2}, \frac{1}{2}] \\ RC & = & [1, -1, -1] \\ RD & = & \frac{1}{2} \end{array}$$

The above examples will be used as running examples in the paper.

Definition 2 *(***Implication***) A system of LDEs $CX = D$* **implies** *a (single) LDE $AX = B$, if every integral vector X satisfying $CX = D$ also satisfies $AX = B$.*

Similarly, $CX = D$ **implies** *a (single) LME $AX \equiv_m B$, if every integral vector X satisfying $CX = D$ also satisfies $AX \equiv_m B$.*

Lemma 1 *(Linear combination) For every rational row vector U the system of LDEs $CX = D$ implies the LDE $UCX = UD$. Note that $UCX = UD$ is simply a linear combination of the equations in $CX = D$. The system $CX = D$ also implies the LME $UCX \equiv_m UD$ for any rational number m.*

Example 4 The system of LDEs $CX = D$ in Example 3 implies the LDE $[\frac{1}{2}, \frac{1}{2}]CX = [\frac{1}{2}, \frac{1}{2}]D$, which simplifies to $x - y - z = \frac{1}{2}$. The system $CX = D$ also implies the LME $x - y - z \equiv_m \frac{1}{2}$ for any rational number m.

3.1 Computing interpolants for systems of LDEs

Let $F \wedge G$ denote an unsatisfiable system of LDEs. The following example shows that an unsatisfiable system of LDEs does not always have an LDE as an interpolant.

Example 5 Let $F := x - 2y = 0$ and $G := x - 2z = 1$. Intuitively, F expresses the constraint that x is even and G expresses the constraint that x is odd, thus, $F \wedge G$ is unsatisfiable. We gave a proof of unsatisfiability of $F \wedge G$ in Example 3. Observe that the pair (F, G) does not have any quantifier-free interpolant that is also a LDE. The problem is that the interpolant can only refer to the variable x. We can prove (using Lemma 6 or see Appendix A) that there is no formula I of the form $c_1 x + c_2 = 0$, where c_1, c_2 are rational numbers, such that $F \Rightarrow I$ and $I \wedge G$ is unsatisfiable.

As shown by the above example it is possible that there exists no LDE that is an interpolant for (F, G). We show that in this case the system (F, G) always has an LME as an interpolant. In the above example an interpolant will be $x \equiv_2 0$. Intuitively, the interpolant means that x is an even integer.

We now describe the algorithm for obtaining interpolants. Let $AX = A', BX = B'$ be systems of LDEs, where $X = [x_1, \ldots, x_n]$ is a column vector of n integer variables. Suppose the combined system of LDEs $AX = A' \wedge BX = B'$ is unsatisfiable. We want to compute an interpolant for $(AX = A', BX = B')$. Let $R = [R_1, R_2]$ be a proof of unsatisfiability of $AX = A' \wedge BX = B'$ according to definition 1. Then

$$R_1 A + R_2 B \quad \text{is integral and} \quad R_1 A' + R_2 B' \quad \text{is not an integer.}$$

Recall that a variable is said to *occur* in a system of LDEs if it occurs with a non-zero coefficient in one of the equations in the system of LDEs. Let $V_{AB} \subseteq X$ denote the set of variables that occur in both $AX = A'$ and $BX = B'$, let $V_{A \setminus B} \subseteq X$ denote the set of variables occurring only in $AX = A'$ (and not in $BX = B'$), and let $V_{B \setminus A} \subseteq X$ denote the set of variables occurring only in $BX = B'$ (and not in $AX = A'$).

We call the LDE $R_1AX = R_1A'$ a **partial interpolant** for $(AX = A', BX = B')$. It is a linear combination of equations in $AX = A'$. The partial interpolant $R_1AX = R_1A'$ can be written in the following form

$$\sum_{x_i \in V_{A\setminus B}} a_i x_i + \sum_{x_i \in V_{AB}} b_i x_i = c \tag{1}$$

where all coefficients a_i, b_i and $c = R_1A'$ are rational numbers. Observe that the partial interpolant does not contain any variable that occurs only in $BX = B'$ ($V_{B\setminus A}$).

Lemma 2 *The coefficient a_i of each $x_i \in V_{A\setminus B}$ in the partial interpolant $R_1AX = R_1A'$ (Equation 1) is an integer.*

Lemma 3 *The partial interpolant $R_1AX = R_1A'$ satisfies the first two conditions in the definition of an interpolant. That is,*
1. $AX = A'$ implies $R_1AX = R_1A'$
2. $(R_1AX = R_1A') \wedge BX = B'$ is unsatisfiable
If $a_i = 0$ for all $x_i \in V_{A\setminus B}$ (equation 1), then the partial interpolant only contains the variables from V_{AB}. In this case the partial interpolant is an interpolant for $(AX = A', BX = B')$.

The proof of above lemmas are given in the appendix A.

Example 6 Consider the system of LDEs $CX = D$ in Example 2. A proof of unsatisfiability for this system is $R = [\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}]$. Let $AX = A'$ be the first two equations in $CX = D$, that is, $x + y = 1 \wedge x - y = 1$ (in matrix form). Let $BX = B'$ be the third equation in $CX = D$, that is, $2y + 2z = 3$. Observe that $V_{A\setminus B} := \{x\}, V_{AB} := \{y\}, V_{B\setminus A} := \{z\}$. In this case $R_1 = [\frac{1}{2}, -\frac{1}{2}]$. The partial interpolant for the pair $(AX = A', BX = B')$ is $y = 0$, which is also an interpolant because $y \in V_{AB}$.

The following example shows that a partial interpolant need not be an interpolant.

Example 7 Consider the system $CX = D$ in Example 3. A proof of unsatisfiability for this system is $R = [\frac{1}{2}, \frac{1}{2}]$. Let $AX = A'$ be the first equation in $CX = D$, that is, $x - 2y = 0$. Let $BX = B'$ be the second equation in $CX = D$, that is, $x - 2z = 1$. Observe that $V_{A\setminus B} := \{y\}, V_{AB} := \{x\}, V_{B\setminus A} := \{z\}$. In this case $R_1 = [\frac{1}{2}]$. Thus, the partial interpolant for the pair $(AX = A', BX = B')$ is $\frac{1}{2}x - y = 0$. Observe that the partial interpolant is not an interpolant as it contains the variable y, which does not occur in V_{AB}. This is not surprising since we have already seen in Example 5 that $(x - 2y = 0, x - 2z = 1)$ cannot have an interpolant that is a LDE.

We now intuitively describe how to remove variables from the partial interpolant that are not common to $AX = A'$ and $BX = B'$. In example 7 the partial interpolant is $\frac{1}{2}x - y = 0$, where $y \notin V_{AB}$. We show how to eliminate y from $\frac{1}{2}x - y = 0$ in order to obtain an interpolant. We use modular arithmetic in order to eliminate y. Informally, the equation $\frac{1}{2}x - y = 0$ implies $\frac{1}{2}x - y \equiv 0 \ (mod\ \gamma)$ for any rational number γ. Let α denote the greatest common divisor of the coefficients of variables (in $\frac{1}{2}x - y = 0$) that do not occur in V_{AB}. In this example $\alpha = 1$ (gcd of the coefficient of y). We know $\frac{1}{2}x - y = 0$ implies $\frac{1}{2}x - y \equiv 0 \ (mod\ 1)$. Since y is an integer variable $y \equiv 0 \ (mod\ 1)$. We can add $\frac{1}{2}x - y \equiv 0 \ (mod\ 1)$ and $y \equiv 0 \ (mod\ 1)$ to obtain $\frac{1}{2}x \equiv 0 \ (mod\ 1)$ (note that y is eliminated). Intuitively, the linear modular equation $\frac{1}{2}x \equiv 0 \ (mod\ 1)$ is an interpolant for $(x - 2y = 0, x - 2z = 1)$. By using basic modular arithmetic this interpolant can be written as $x \equiv 0 \ (mod\ 2)$.

We now formalize the above intuition to address the case when the partial interpolant contains variables that are not common to $AX = A'$ and $BX = B'$.

Theorem 2 *Assume that the coefficient a_i of at least one $x_i \in V_{A\setminus B}$ in the partial interpolant (Equation 1) is not zero. Let α denote the gcd of $\{a_i | x_i \in V_{A\setminus B}\}$.*
(a) α is an integer and $\alpha > 0$.
(b) Let β be any integer that divides α. Then the following linear modular equation I_β is an interpolant for $(AX = A', BX = B')$.

$$I_\beta := \sum_{x_i \in V_{AB}} b_i x_i \equiv c \; (mod \; \beta)$$

Observe that I_β contains only variables that are common to both $AX = A'$ and $BX = B'$. It is obtained from the partial interpolant by dropping all variables occurring only in $AX = A'$ ($V_{A\setminus B}$) and replacing the linear equality by a modular equality.

The proof can be found in the appendix A.2. In theorem 2, I_1 is always an interpolant for $(AX = A', BX = B')$. For $\alpha > 1$ theorem 2 allows us to obtain multiple interpolants by choosing different β. For any β that divides α, $I_\alpha \Rightarrow I_\beta$ and $I_\beta \Rightarrow I_1$. Depending upon the application one can use the strongest interpolant I_α (least satisfying assignments) or the weakest interpolant I_1 (most satisfying assignments). The next example illustrates the use of Theorem 2 in obtaining multiple interpolants.

Example 8 Consider the system of LDEs $CX = D$ and a proof of unsatisfiability R:

$$CX = D := \begin{bmatrix} 30 & 4 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix} \qquad \begin{array}{lcl} R & = & [\frac{1}{5}, \frac{1}{5}] \\ RC & = & [6, 1] \\ RD & = & \frac{4}{5} \end{array}$$

Let $AX = A'$ be the first equation in $CX = D$, that is, $30x + 4y = 2$ (in matrix form). Let $BX = B'$ be the second equation in $CX = D$, that is, $y = 2$. Observe that $V_{A\setminus B} := \{x\}, V_{AB} := \{y\}, V_{B\setminus A} := \emptyset$. In this case $R_1 = [\frac{1}{5}]$. The partial interpolant $R_1 AX = R_1 A'$ for the pair $(AX = A', BX = B')$ is $6x + \frac{4}{5}y = \frac{2}{5}$. The partial interpolant is not an interpolant as it contains the variable x, which does not occur in V_{AB}.

Using Theorem 2 we can obtain four interpolants for the pair $(AX = A', BX = B')$:

$$\begin{array}{lcl} I_1 & := & \frac{4}{5}y \equiv_1 \frac{2}{5} \\ I_2 & := & \frac{4}{5}y \equiv_2 \frac{2}{5} \\ I_3 & := & \frac{4}{5}y \equiv_3 \frac{2}{5} \\ I_6 & := & \frac{4}{5}y \equiv_6 \frac{2}{5} \end{array}$$

I_6 implies all other interpolants. That is, $I_6 \Rightarrow I_3, I_6 \Rightarrow I_2, I_6 \Rightarrow I_1$. I_1 is implied by all other interpolants. That is, $I_2 \Rightarrow I_1, I_3 \Rightarrow I_1, I_6 \Rightarrow I_1$.

Lemma 3 and Theorem 2 give us a sound and complete algorithm for computing an interpolant for unsatisfiable systems of LDEs. (See Appendix A.3 for the algorithm pseudocode.)

4 System of linear modular equations (LMEs)

In this section we discuss proofs of unsatisfiability and interpolation algorithm for LMEs. We first consider a system of LMEs where all equations have the same modulus l, where l is a rational number. We denote this

system as $CX \equiv_l D$, where C denotes an $m \times n$ rational matrix, X denotes a column vector of n integer variables and D denotes a column vector of m rational numbers. The next theorem gives a necessary and sufficient condition for $CX \equiv_l D$ to have an integral solution.

Theorem 3 *The system $CX \equiv_l D$ has no integral solution X if and only if there exists a rational row vector R such that RC is integral, lR is integral, and RD is not an integer. Note that lR denotes the row vector obtained by multiplying each element of R by rational number l. (The size of R is $1 \times m$.)*

The proof uses reduction to LDEs. See the appendix B.1 for the proof.

Definition 3 *We say a system of LMEs $CX \equiv_l D$ is* **unsatisfiable** *if it has no integral solution X. A* **proof of unsatisfiability** *for a system of LMEs $CX \equiv_l D$ is a rational row vector R such that RC is integral, lR is integral, and RD is not an integer.*

Example 9 Consider the system of LMEs $CX \equiv_8 D$ and a proof of unsatisfiability R:

$$CX \equiv_8 D := \begin{bmatrix} 2 & 2 \\ 2 & 1 \\ 4 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \equiv_8 \begin{bmatrix} 4 \\ 4 \\ 4 \end{bmatrix} \qquad \begin{array}{lcl} R & = & [\frac{1}{4}, -\frac{1}{2}, -\frac{1}{8}] \\ RC & = & [-1, 0] \\ lR & = & [2, -4, -1] \\ RD & = & -\frac{3}{2} \end{array}$$

Intuitively, $CX \equiv_8 D$ is unsatisfiable because we can take an integer linear combination of the given equations using lR to get a contradiction $0 \equiv_8 -12$.

Definition 4 *(***Implication***) A system of LMEs $CX \equiv_l D$* **implies** *a LME $AX \equiv_l B$, if every integral vector X satisfying $CX \equiv_l D$ also satisfies $AX \equiv_l B$.*

Lemma 4 *For every* **integral** *row vector U the system of LMEs $CX \equiv_l D$ imply $UCX \equiv_l UD$.*

4.1 Computing interpolants for systems of LMEs

Let $AX \equiv_l A'$ and $BX \equiv_l B'$ be two systems of LMEs such that $AX \equiv_l A' \wedge BX \equiv_l B'$ is unsatisfiable. We show that $(AX \equiv_l A', BX \equiv_l B')$ always has an LME as an interpolant. Let $R = [R_1, R_2]$ denote a proof of unsatisfiability for the system $AX \equiv_l A' \wedge BX \equiv_l B'$ such that $R_1A + R_2B$ is integral, $lR = [lR_1, lR_2]$ is integral, and $R_1A' + R_2B'$ is not an integer. The following theorem shows that we can take integer linear combinations of equations in $AX \equiv_l A'$ to obtain interpolants.

Theorem 4 *We assume $l \neq 0$. Let S_1 denote the set of non-zero coefficients of $x_i \in V_{A \setminus B}$ in R_1AX. Let S_2 denote the set of non-zero elements of row vector lR_1. If $S_2 = \emptyset$, then the interpolant for $(AX \equiv_l A', BX \equiv_l B')$ is a trivial LME $0 \equiv_l 0$. Otherwise, let $S_2 \neq \emptyset$. Let α denote the gcd of numbers in $S_1 \cup S_2$.*
(a) α is an integer and $\alpha > 0$.
(b) Let β be any integer that divides α. Let $U = \frac{l}{\beta} R_1$. Then $UAX \equiv_l UA'$ is an interpolant for $(AX \equiv_l A', BX \equiv_l B')$.

The proof is given in the appendix B.2.

Example 10 Consider the system of LMEs $CX \equiv_l D$ in Example 9. Let $AX \equiv_l A'$ denote the first two equations in $CX \equiv_l D$ and $BX \equiv_l B'$ denote the last equation in $CX \equiv_l D$. Observe that $V_{A \setminus B} := \{y\}, V_{AB} := \{x\}, V_{B \setminus A} := \emptyset$. A proof of unsatisfiability for $CX \equiv_l D$ is $R = [\frac{1}{4}, -\frac{1}{2}, -\frac{1}{8}]$. We have $R_1 = [\frac{1}{4}, -\frac{1}{2}]$, $lR_1 = [2, -4]$, R_1AX is $-\frac{1}{2}x$, $S_1 = \emptyset$, $S_2 = \{2, -4\}$, $\alpha = 2$. We can take $\beta = 1$ or $\beta = 2$ to obtain two valid interpolants. For $\beta = 1$, $U = [2, -4]$ and the interpolant $UAX \equiv_l UA'$ is $-4x \equiv_8 -8$ (equivalently $x \equiv_2 0$). For $\beta = 2$, $U = [1, -2]$ and the interpolant $UAX \equiv_l UA'$ is $-2x \equiv_8 -4$ (equivalently $x \equiv_4 2$).

4.2 Handling LMEs with different moduli

Consider a system F of LMEs, where equations in F can have different moduli. In order to check the satisfiability of F, we obtain another equivalent system of equations F' such that each equation in F' has the same moduli. This is done using a standard trick described in Mathews [22]. Let $m_1, \ldots, m_k$ represent the different moduli occurring in equations in F. Let m denote the least common multiple of $m_1, \ldots, m_k$. We multiply each equation $t \equiv_{m_i} c$ in F by $\frac{m}{m_i}$ to obtain another equation $\frac{m}{m_i}t \equiv_m \frac{m}{m_i}c$. Let F' represent the set of new equations. All equations in F' have same modulus m. Using basic modular arithmetic one can show that F and F' are equivalent. Suppose F is unsatisfiable. Then the interpolants for any partition of F can be computed by working with F' and using the techniques described in the previous section. For example, let F represent the following system of LMEs $x \equiv_2 1 \wedge x + y \equiv_4 2 \wedge 2x + y \equiv_8 4$. One can work with $F' := 4x \equiv_8 4 \wedge 2x + 2y \equiv_8 4 \wedge 2x + y \equiv_8 4$ instead of F.

5 Algorithms for obtaining Proofs of Unsatisfiability

Polynomial time algorithms are known for determining if a system of LDEs $CX = D$ has an integral solution or not [29]. We review one such algorithm that is based on the computation of the *Hermite normal form (HNF)* of the matrix C.

Using standard Gaussian elimination it can be determined if $CX = D$ has a rational solution or not. If $CX = D$ has no rational solution, then it cannot have any integral solution. In the discussion below we assume that $CX = D$ has a rational solution. Without loss of generality we assume that matrix C has *full row rank*, that is, all rows of C are linearly independent (linearly dependent equations can be removed).

The HNF of a $m \times n$ matrix C with full row rank is of the form $[E\ \ 0]$ where 0 represents an $m \times (n-m)$ matrix filled with zeros and E is a square $m \times m$ matrix with the following properties: 1) E is lower triangular 2) E is non-singular (invertible) 3) all entries in E are non-negative and the maximum entry in each row lies on the diagonal. The HNF of a matrix can be obtained by three elementary column operations. 1) Exchanging two columns. 2) Multiplying a column by -1. 3) Adding an integral multiple of one column to another column. Each column operation can be represented by a unimodular matrix. A *unimodular matrix* is a square matrix with integer entries and determinant +1 or -1. The product of unimodular matrices is a unimodular matrix. The inverse of a unimodular matrix is a unimodular matrix. The conversion of C to HNF can be represented as follows $CU = [E\ \ 0]$, where U is a unimodular matrix, the sizes of C, U, E are $m \times n, n \times n, m \times m$, respectively and 0 represents an $m \times (n - m)$ matrix filled with zeros ($n \geq m$ because C has full row-rank). The following result shows the use of HNF in determining the satisfiability of a system of LDEs. Let E^{-1} denotes the matrix inverse of E.

Lemma 5 *(Schrijver [29]) For C, X, D, E defined as above, $CX = D$ has no integral solution if and only if $E^{-1}D$ is not integral.*

Example 11 For the system of LDEs $CX = D$ in example 2 we have the following:

$$\underbrace{\begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 2 & 2 \end{bmatrix}}_{C} \underbrace{\begin{bmatrix} 1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 1 & 1 \end{bmatrix}}_{U} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}}_{E} \quad \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ \frac{-1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} \end{bmatrix}}_{E^{-1}} \underbrace{\begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix}}_{D} = \underbrace{\begin{bmatrix} 1 \\ 0 \\ \frac{3}{2} \end{bmatrix}}_{not\ integral}$$

Example 12 For the system of LDEs $CX = D$ in example 3 we have the following:

$$\underbrace{\begin{bmatrix} 1 & -2 & 0 \\ 1 & 0 & -2 \end{bmatrix}}_{C} \underbrace{\begin{bmatrix} 1 & 2 & -2 \\ 0 & 1 & -1 \\ 0 & 0 & -1 \end{bmatrix}}_{U} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \end{bmatrix}}_{[E\ 0]} \quad \underbrace{\begin{bmatrix} 1 & 0 \\ \frac{-1}{2} & \frac{1}{2} \end{bmatrix}}_{E^{-1}} \underbrace{\begin{bmatrix} 0 \\ 1 \end{bmatrix}}_{D} = \underbrace{\begin{bmatrix} 0 \\ \frac{1}{2} \end{bmatrix}}_{not\ integral}$$

5.1 Obtaining a proof of unsatisfiability for a system of LDEs

If a system of LDEs $CX = D$ is unsatisfiable, then we want to compute a row vector R such that RC is integral and RD is not an integer. The following corollary shows that the proof of unsatisfiability can be obtained by using the HNF of C.

Corollary 1 *Given $CX = D$ where C, D are rational matrices, and C has full row rank. Let $[E\ 0]$ denote the HNF of C. If $CX = D$ has no integral solution, then $E^{-1}D$ is not integral. Suppose the i^{th} entry in $E^{-1}D$ is not an integer. Let R' denote the i^{th} row in E^{-1}. Then (a) $R'D$ is not an integer and (b) $R'C$ is integral. Thus, R' serves as the required proof of unsatisfiability of $CX = D$.*

The proof is given in the appendix C.

Example 13 In example 11 the third row in $E^{-1}D$ is not an integer. Thus, the proof of unsatisfiability of $CX = D$ is the third row in E^{-1} which is $[0, 0, \frac{1}{2}]$.

In example 12 the second row in $E^{-1}D$ is not an integer. Thus, the proof of unsatisfiability of $CX = D$ is the second row in E^{-1} which is $[-\frac{1}{2}, \frac{1}{2}]$.

Proofs of unsatisfiability for LMEs Let $CX \equiv_l D$ be a system of LMEs. Each equation $t_i \equiv_l d_i$ in $CX \equiv_l D$ can be written as an equi-satisfiable LDE, $t_i + lv_i = d_i$, where v_i is a new integer variable. In this way we can reduce the given $CX \equiv_l D$ to an equi-satisfiable system of LDEs $C'Z = D$. The proof of unsatisfiability of $C'Z = D$ is exactly a proof of unsatisfiability of $CX \equiv_l D$ (see the proof of theorem 3).

Complexity If a system of LDEs or LMEs is unsatisfiable, then we can obtain a proof of unsatisfiability in polynomial time. This is because HNF computation, matrix inversion, and matrix multiplication can be done in polynomial time in the size of input [29, 31]. The interpolation algorithms described in Sections 3 and 4 are polynomial in the size of the given formulas and the proof of unsatisfiability.

6 Handling Linear Diophantine Equations and Disequations

We show how to compute interpolants in presence of linear diophantine disequations. A *linear diophantine disequation (LDD)* is of the form $c_1x_1 + \ldots + c_nx_n \neq c_0$, where $c_0, \ldots, c_n$ are rational numbers and $x_1, \ldots, x_n$ are integer variables. *A system of LDEs+LDDs* denotes conjunctions of LDEs and LDDs. For example, $x + 2y = 1 \wedge x + y \neq 1 \wedge 2y + z \neq 1$ with x, y, z as integer variables represents a system of LDEs+LDDs. We represent a conjunction of m LDDs as $\bigwedge_{i=1}^{m} C_iX \neq D_i$, where C_i is a rational row vector and D_i is a rational number. The next theorem gives a necessary and sufficient condition for a system of LDEs+LDDs to have an integral solution.

Theorem 5 *Let F denote $AX = B \wedge \bigwedge_{i=1}^{m} C_iX \neq D_i$. The following are equivalent:*
1. F has no integral solution
2. F has no rational solution or $AX = B$ has no integral solution.

The proof of $(2) \Rightarrow (1)$ in Theorem 5 is easy. The proof of $(1) \Rightarrow (2)$ is involved and relies on the following lemmas (full proof is given in the appendix F). The first lemma shows that if a system of LDEs $AX = B$ has an integral solution, then every LDE that is implied by $AX = B$, can be obtained by a linear combination of equations in $AX = B$.

Lemma 6 *A system of LDEs $AX = B$ implies a LDE $EX = F$ if and only if $AX = B$ is unsatisfiable or there exists a rational vector R such that $E = RA$ and $F = RB$.*

We use the properties of the *cutting-plane* proof system [29, 7] in order to prove lemma 6. The proof is given in the appendix D. The next lemma shows that if a system of LDEs implies a disjunction of LDEs, then it implies one of the LDEs in the disjunction (also called *convexity* [25]).

Lemma 7 *A system of LDEs $AX = B$ implies $\bigvee_{i=1}^{m} C_i X = D_i$ if and only if there exists $1 \leq k \leq m$ such that $AX = B$ implies $C_k X = D_k$.*

We use a theorem from [29] that gives a parametric description of the integral solutions to $AX = B$ in order to prove lemma 7. See the appendix E for the full proof. Let F denote $AX = B \wedge \bigwedge_{i=1}^{m} C_i X \neq D_i$. Using Theorem 5 we can determine whether F has an integral solution in polynomial time. This is because checking if $AX = B$ has an integral solution can be done in polynomial time [29, 7]. Checking whether the system F has a rational solution can be done in polynomial time as well [25].

6.1 Interpolants for LDEs+LDDs

We say a system of LDEs+LDDs is **unsatisfiable** if it has no integral solution. Consider systems of LDEs+LDDs $F := F_1 \wedge F_2$ and $G := G_1 \wedge G_2$, where F_1, G_1 are systems of LDEs and F_2, G_2 are systems of LDDs. $F \wedge G$ represents another system of LDEs+LDDs. Suppose $F \wedge G$ is unsatisfiable. The interpolant for (F, G) can be computed by considering two cases (due to theorem 5):
Case 1: $F \wedge G$ is unsatisfiable because $F_1 \wedge F_2 \wedge G_1 \wedge G_2$ has no rational solution. We can compute an interpolant for (F, G) using the techniques described in [24, 33, 28, 10]. For completeness we describe this case in the appendix G. The interpolant can be a LDE or a LDD.
Case 2: $F \wedge G$ is unsatisfiable because $F_1 \wedge G_1$ has no integral solution. In this case we can compute an interpolant for the pair (F_1, G_1) using the techniques from Section 3. The interpolant for (F_1, G_1) will be an interpolant for (F, G). It can be a LDE or a LME.

7 Experimental results

We implemented the interpolation algorithms for conjunctions of LDEs, LMEs, LDDs in a tool called `INT2 (INTeger INTerpolate)`. The experiments are performed on a 1.86 GHz Intel Xeon (R) machine with 4 GB of memory running Linux. `INT2` is designed for computing interpolants for formulas (LDEs, LMEs, LDEs+LDDs) that are satisfiable over rationals but unsatisfiable over integers. Currently, there are no other interpolation tools for such formulas.

7.1 Use of Interpolants in Verification

We wrote a collection of small C programs each containing a `while` loop and an ERROR label. These programs are safe (ERROR is unreachable). The existing tools based on predicate abstraction and counterexample guided abstraction refinement (CEGAR) such as `BLAST` [1, 17], `SATABS` [2] are not able to

Example	Preds/Interpolants	VINT2
ex1	$y \equiv_2 1$	2.72s
ex2	$x + y \equiv_2 0$	0.83s
ex4	$x + y + z \equiv_4 0$	0.95s
ex5	$x \equiv_4 0, y \equiv_4 0$	1.1s
ex6	$4x + 2y + z \equiv_8 0$	0.93s
ex7	$4x - 2y + z \equiv_{2^{22}} 0$	0.54s
forb1	$x + y \equiv_3 0$	-

Table 1: Table showing the predicates needed and time taken in seconds.

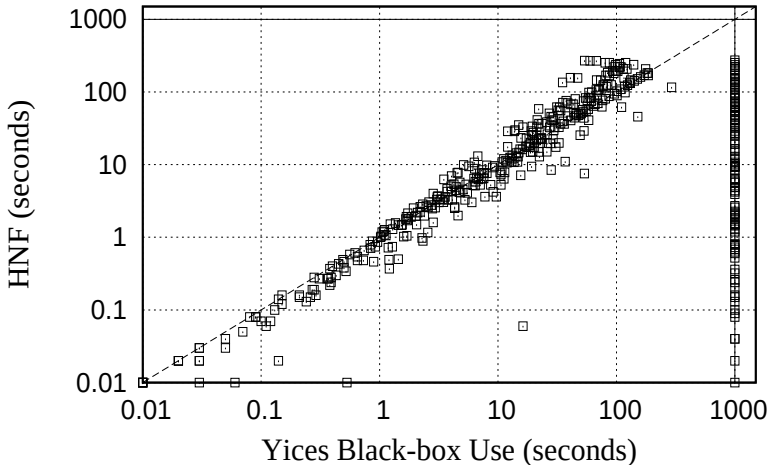

Figure 1: Comparing Hermite Normal Form based algorithm and black-box use of Yices for getting proofs of unsatisfiability

verify these programs. This is because the inductive invariant required for the proof contains LMEs as predicates, shown in the "Preds/Interpolants" column of Table 1. These predicates cannot be discovered by the interpolation engine [24, 28] used in `BLAST` or by the weakest precondition based procedure used in `SATABS`. The interpolation algorithms described in this paper are able to find the right predicates by computing the interpolants for spurious program traces. Only one unwinding of the `while` loop suffices to find the right predicates in 6 out of 7 cases. In program ex5 multiple unwindings of the `while` loop produces predicates of the form $x = 0, y = 4, x = 4, y = 8, \ldots$. After a few unwindings these predicates are generalized to obtain $x \equiv_4 0, y \equiv_4 0$ (by taking gcd of the numbers involved).

We wrote similar programs in Verilog and tried verifying them with `VCEGAR` [3], a CEGAR based model checker for Verilog. `VCEGAR` fails on these examples due to its use of weakest preconditions. Next, we externally provided the interpolants (predicates) found by `INT2` to `VCEGAR`. With the help of these predicates `VCEGAR` is able to show the unreachability of ERROR labels in all examples except forb1 (ERROR is reachable in the Verilog version of forb1). The runtimes are shown in "VINT2" column.

7.2 Proofs of unsatisfiability (PoU) algorithms

We obtained 459 unsatisfiable formulas (system of LDEs) by unwinding the `while` loops for C programs mentioned above. The number of LDEs in these formulas range from 3 to 1500 with 2 to 4 variables per equation. There are two options for obtaining PoU in INT2.

(a) Using Hermite Normal Form (HNF) (Section 5.1). We use PARI/GP [32] to compute HNF of matrices.

(b) By using a state-of-the-art SMT solver Yices 1.0.11 [4] in a black-box fashion (along the lines of [28]). Given a system of LDEs $AX = B$ we encode the constraints that RA is integral and RB is not an integer by means of mixed integer linear arithmetic constraints (see the appendix J). The SMT solver returns concrete values to elements in R if $AX = B$ is unsatisfiable.

The comparison between (a) and (b) is shown in Figure 1. There is a timeout of 1000 seconds per problem. The HNF based algorithm is able to solve all problems, while the black-box usage of Yices cannot solve 102 problems within the timeout. Thus, the HNF based method is superior over the black-box use of Yices.

We also ran Yices to decide whether $AX = B$ has an integral solution or not. The system $AX = B$ (X integral) is given to Yices. In this case, Yices is very efficient and reports the satisfiability or unsatisfiability of $AX = B$ quickly. However, no PoU is provided when $AX = B$ is unsatisfiable. In principle it is possible for Yices to provide a PoU when $AX = B$ is unsatisfiable (although this will add some overhead).

Note that the interpolation algorithms proposed in our paper are independent of the algorithm used to generate the PoU. Any decision procedure that can produce PoU according to definitions 1, 3 can be used (we are not restricted to using HNF or Yices).

8 Conclusion

We presented polynomial time algorithms for computing proofs of unsatisfiability and interpolants for conjunctions of linear diophantine equations, linear modular equations and linear diophantine disequations. These interpolation algorithms are useful for discovering modular/divisibility predicates from spurious counterexamples in a counterexample guided abstraction refinement framework. In future, we plan to work on interpolating theorem provers for integer linear arithmetic and bit-vector arithmetic and make use of the satisfiability modulo theories framework.

Acknowledgment. We thank Axel Legay and Jeremy Avigad for their valuable comments.

References

[1] BLAST 2.4 website. `http://mtc.epfl.ch/software-tools/blast/`.

[2] SATABS 1.9 website, `http://www.verify.ethz.ch/satabs/`.

[3] VCEGAR 1.3 website. `http://www.cs.cmu.edu/~modelcheck/vcegar/`.

[4] Yices 1.0.11 website. `http://yices.csl.sri.com/`.

[5] Domagoj Babić and Madanlal Musuvathi. Modular Arithmetic Decision Procedure. Technical Report TR-2005-114, Microsoft Research Redmond, 2005.

[6] Clark W. Barrett, David L. Dill, and Jeremy R. Levitt. A decision procedure for bit-vector arithmetic. In *DAC '98: Proceedings of the 35th annual conference on Design automation*, pages 522–527, New York, NY, USA, 1998. ACM Press.

[7] Alexander Bockmayr and Volker Weispfenning. Solving numerical constraints. In A. Robinson and A. Voronkov, editors, *Handbook of Automated Reasoning*, pages 751–842. 2001.

[8] Roberto Bruttomesso, Alessandro Cimatti, Anders Franzen, Alberto Griggio, Ziyad Hanna, Alexander Nadel, Amit Palti, and Roberto Sebastiani. A lazy and layered smt(bv) solver for hard industrial verification problems. In *Computer Aided Verification (CAV '07)*, Berlin, Germany, July 2007. Springer-Verlag.

[9] R. E. Bryant, D. Kroening, J. Ouaknine, S. A. Seshia, O. Strichman, and B. Brady. Deciding bit-vector arithmetic with abstraction. In *Tools and Algorithms for the Construction and Analysis of Systems (TACAS)*, 2007.

[10] Alessandro Cimatti, Alberto Griggio, and Roberto Sebastiani. Efficient interpolation in satisfiability modulo theories. In *TACAS*, 2008. To appear.

[11] E. Clarke, O. Grumberg, S. Jha, Y. Lu, and H. Veith. Counterexample-guided abstraction refinement for symbolic model checking. *J. ACM*, 50(5), 2003.

[12] William Craig. Linear reasoning. a new form of the herbrand-gentzen theorem. *J. Symb. Log.*, 22(3):250–268, 1957.

[13] David Cyrluk, M. Oliver Möller, and Harald Rueß. An efficient decision procedure for the theory of fixed-sized bit-vectors. In *CAV '97: Proceedings of the 9th International Conference on Computer Aided Verification*, pages 60–71, London, UK, 1997. Springer-Verlag.

[14] Bruno Dutertre and Leonardo Mendonça de Moura. A fast linear-arithmetic solver for dpll(t). In *CAV*, pages 81–94, 2006.

[15] Vijay Ganesh, Sergey Berezin, and David L. Dill. A decision procedure for fixed-width bit-vectors. Technical Report CSTR 2007-06, Stanford Computer Science Department, 2005.

[16] Vijay Ganesh and David L. Dill. A decision procedure for bit-vectors and arrays. In *Computer Aided Verification (CAV '07)*, Berlin, Germany, July 2007. Springer-Verlag.

[17] Thomas A. Henzinger, Ranjit Jhala, Rupak Majumdar, and Kenneth L. McMillan. Abstractions from proofs. In *POPL*, pages 232–244. ACM Press, 2004.

[18] Ranjit Jhala and Kenneth L. McMillan. A practical and complete approach to predicate refinement. In *TACAS*, pages 459–473, 2006.

[19] Deepak Kapur, Rupak Majumdar, and Calogero G. Zarba. Interpolation for data structures. In *SIGSOFT '06/FSE-14*, pages 105–116. ACM, 2006.

[20] Daniel Kroening and Georg Weissenbacher. Lifting propositional interpolants to the word-level. In *FMCAD*, pages 85–89. IEEE, 2007.

[21] P. Manolios, S. K. Srinivasan, , and D. Vroon. Automatic memory reductions for RTL-level verification. In *ICCAD*, 2006.

[22] George Ballard Mathews. *Theory of numbers*. NY, Chelsea Pub. Co., 2nd edition, 1927.

[23] K. L. McMillan. Interpolation and sat-based model checking. In *CAV*, pages 1–13, 2003.

[24] K. L. McMillan. An interpolating theorem prover. In *TACAS*, pages 16–30, 2004.

[25] Greg Nelson and Derek C. Oppen. Simplification by cooperating decision procedures. *ACM Trans. Program. Lang. Syst.*, 1(2):245–257, 1979.

[26] Mojżesz Presburger. Über die vollständigkeit eines gewissen systems der arithmetik ganzer zahlen, in welchem die addition als einzige operation hervortritt. In *Sprawozdanie z I Kongresu metematyków slowiańskich, Warszawa 1929*, pages 92–101,395, Warsaw, Poland, 1930. Annotated English version in [30].

[27] Pavel Pudlák. Lower bounds for resolution and cutting plane proofs and monotone computations. *J. Symb. Log.*, 62(3):981–998, 1997.

[28] Andrey Rybalchenko and Viorica Sofronie-Stokkermans. Constraint solving for interpolation. In *VMCAI*, pages 346–362, 2007.

[29] A. Schrijver. *Theory of linear and integer programming*. John Wiley & Sons, NY, 1986.

[30] R. Stansifer. Presburger's article on integer arithmetic: Remarks and translation. Technical Report TR84–639, Cornell University Computer Science Department, 1984.

[31] Arne Storjohann and George Labahn. Asymptotically fast computation of hermite normal forms of integer matrices. In *ISSAC '96: Proceedings of the 1996 international symposium on Symbolic and algebraic computation*, pages 259–266, 1996.

[32] The PARI Group. *PARI/GP, Version 2.3.2*, 2006. `http://pari.math.u-bordeaux.fr/`.

[33] Greta Yorsh and Madanlal Musuvathi. A combination method for generating interpolants. In *CADE*, pages 353–368, 2005.

A Proofs from Section 3

Proof of Lemma 1

Proof. $UCX = UD$ is a linear combination of equations in $CX = D$. Let X_0 be an integral solution to $CX = D$. It is easy to verify that X_0 also satisfies $UCX = UD$. Thus, the system of LDEs $CX = D$ implies the LDE $UCX = UD$ for any rational row vector U.

Since $UCX_0 - UD = 0$, any rational number m divides $UCX_0 - UD$. It follows that X_0 is also a solution to the LME $UCX \equiv_m UD$. Thus, the system of LDEs $CX = D$ implies the LME $UCX \equiv_m UD$ for any rational row vector U and rational number m. □

Why $F \wedge G$ has no LDE as interpolant in Example 5.

Proof. Recall, that F is $x - 2y = 0$ and G is $x - 2z = 1$, where x, y, z are integers. Observe that F has an integral solution, for example, $x = 2, y = 1$. Thus, by lemma 6 any LDE that is implied by F must be of the form $r(x - 2y = 0)$, where r is a rational number.

Suppose (F, G) have an LDE I as an interpolant. Since $F \Rightarrow I$, I must be of the form $r(x-2y = 0)$. But I can only contain variable x (common variable of F and G). This is possible only when $r = 0$. With $r = 0$, I reduces to $0 = 0$ which is not unsatisfiable with G. Thus, (F, G) cannot have an LDE as an interpolant. □

Proof of Lemma 2

Proof. By definition of $V_{A\setminus B}$ the coefficient of $x_i \in V_{A\setminus B}$ is zero in each equation of $BX = B'$. Thus, the coefficient of $x_i \in V_{A\setminus B}$ must be the same in R_1AX and $(R_1A + R_2B)X$. Since $R_1A + R_2B$ is integral it follows that the coefficient of $x_i \in V_{A\setminus B}$ (a_i) in the partial interpolant is an integer. □

A.1 Proof of Lemma 3

Lemma 3. *The partial interpolant $R_1AX = R_1A'$ satisfies the first two conditions in the definition of an interpolant. That is,*
1. $AX = A'$ implies $R_1AX = R_1A'$
2. $(R_1AX = R_1A') \wedge BX = B'$ is unsatisfiable
If $a_i = 0$ for all $x_i \in V_{A\setminus B}$ (equation 1), then the partial interpolant is also a interpolant for $(AX = B, A'X = B')$. In this case the partial interpolant only contains the variables from V_{AB}.

Proof. 1. $AX = A'$ implies $R_1AX = R_1A'$. This follows from Lemma 1.

2. Observe that $(R_1AX = R_1A') \wedge BX = B'$ is a system of LDEs

$$\begin{bmatrix} R_1A \\ B \end{bmatrix} X = \begin{bmatrix} R_1A' \\ B' \end{bmatrix}$$

We show that the row vector $[1, R_2]$ is a proof of unsatisfiability of $I \wedge (BX = B')$. This requires showing the conditions in the definition of proof of unsatisfiability are met.

- To show

$$[1, R_2] \begin{bmatrix} R_1 A \\ B \end{bmatrix} \text{ is integral.}$$

The above product is equal to $R_1 A + R_2 B$ which is integral.

- To show

$$[1, R_2] \begin{bmatrix} R_1 A' \\ B' \end{bmatrix} \text{ is not an integer.}$$

The above product is equal to $R_1 A' + R_2 B'$ which is not an integer. Thus, $[1, R_2]$ is a proof of unsatisfiability of $I \wedge (BX = B')$. So $I \wedge (BX = B')$ is unsatisfiable. $\square$

A.2 Proof of Theorem 2

Recall that rational row vector $[R_1, R_2]$ is the proof of unsatisfiability of $AX = A' \wedge BX = B'$ (A, B, A', B' are rational matrices) such that

$$\begin{array}{ll} R_1 A + R_2 B & \text{is integral} \\ R_1 A' + R_2 B' & \text{is not an integer} \end{array}$$

We call $R_1 AX = R_1 A'$ the partial interpolant for $(AX = A', BX = B')$. It can be written as follows:

$$\sum_{x_i \in V_{A \setminus B}} a_i x_i + \sum_{x_i \in V_{AB}} b_i x_i = c \tag{2}$$

where all coefficients a_i, b_i and $c = R_1 A'$ are rational numbers. The above equation is the same as Equation 1 repeated here for convenience.

Similarly, $R_2 BX = R_2 B'$ can be written as follows:

$$\sum_{x_i \in V_{AB}} e_i x_i + \sum_{x_i \in V_{B \setminus A}} f_i x_i = d \tag{3}$$

where all coefficients e_i, f_i and $d = R_2 B'$ are rational numbers. Observe that $R_2 BX = R_2 B'$ does not contain any variable from $V_{A \setminus B}$.

Lemma 8 *Using the notation from Equations 2 and 3:*
(a) For all $x_i \in V_{A \setminus B}$, a_i is an integer.
(b) For all $x_i \in V_{AB}$, $b_i + e_i$ is an integer.
(c) For all $x_i \in V_{B \setminus A}$, f_i is an integer.
(d) $c + d$ is not an integer.

Proof. The sum of the left hand sides of Equations 2 and 3 is

$$\sum_{x_i \in V_{A \setminus B}} a_i x_i + \sum_{x_i \in V_{AB}} (b_i + e_i) x_i + \sum_{x_i \in V_{B \setminus A}} f_i x_i$$

which is the same as $(R_1 A + R_2 B)X$. Since $R_1 A + R_2 B$ is integral each coefficient in the above sum must be an integer. This gives us the desired results (a),(b),(c).

Since $c + d = R_1A' + R_2B'$ and $R_1A' + R_2B'$ is not an integer we get (d). $\square$

Theorem 2. *Assume that the coefficient a_i of at least one $x_i \in V_{A\setminus B}$ in the partial interpolant (Equation 2) is not zero. Let α denote the gcd of $\{a_i | x_i \in V_{A\setminus B}\}$.*
(a) α is an integer and $\alpha > 0$.
(b) Let β be any integer that divides α. Then the following linear modular equation I_β is an interpolant for $(AX = A', BX = B')$.

$$I_\beta := \sum_{x_i \in V_{AB}} b_i x_i \equiv c \; (mod \; \beta)$$

Observe that I_β contains only variables that are common to both $AX = A'$ and $BX = B'$. It is obtained from the partial interpolant (Equation 2) by dropping all variables occurring only in $AX = A'$ ($V_{A\setminus B}$) and replacing the linear equality by a modular equality.

Proof. (a) By lemma 8 each a_i is an integer. Since α is the gcd of $\{a_i | x_i \in V_{A\setminus B}\}$, α must be an integer. Also note that α is non-zero since at least one a_i is non-zero. By definition of gcd α is positive.

(b) To show that I_β is an interpolant for $(AX = A', BX = B')$.
1. We need to show that $AX = A'$ implies I_β. Recall, that $AX = A'$ implies the partial interpolant $R_1AX = R_1A'$ from lemma 3. We show that $R_1AX = R_1A'$ implies I_β.

From basic modular arithmetic it follows that $s = t$ implies $s \equiv t \; (mod \; \gamma)$ for any rational number γ. Thus, the partial interpolant $R_1AX = R_1A'$ implies $R_1AX \equiv_\beta R_1A'$, where β is any integer that divides α. Consider the equation form of $R_1AX \equiv_\beta R_1A'$ (equation 2):

$$\sum_{x_i \in V_{A\setminus B}} a_i x_i + \sum_{x_i \in V_{AB}} b_i x_i \equiv_\beta c \tag{4}$$

By definition α divides a_i for all $x_i \in V_{A\setminus B}$. Since β divides α, it follows that β divides a_i for all $x_i \in V_{A\setminus B}$. As x_i is an integer valued variable, $a_i x_i$ is divisible by β for all $x_i \in V_{A\setminus B}$. It follows that

$$\sum_{x_i \in V_{A\setminus B}} a_i x_i \equiv_\beta 0. \tag{5}$$

Subtract equation 5 from equation 4 to obtain

$$\sum_{x_i \in V_{AB}} b_i x_i \equiv_\beta c.$$

The above equation is I_β. $AX = A'$ implies $R_1AX = R_1A'$ and $R_1AX = R_1A'$ implies equation 4. Equation 5 holds for any integral assignment to all $x_i \in V_{A\setminus B}$. So $R_1AX = R_1A'$ implies equation 5. Equations 4, 5 imply I_β. It follows that $AX = A'$ implies I_β.

2. We need to show that $I_\beta \wedge BX = B'$ is unsatisfiable. Assume for the sake of contradiction that $I_\beta \wedge BX = B'$ has an integral satisfying assignment. Let the satisfying assignment to $I_\beta \wedge BX = B'$ be $x_i = g_i$ where g_i is an integer for all $x_i \in V_{AB} \cup V_{B\setminus A}$. Since I_β is satisfied by g_i we have

$$\sum_{x_i \in V_{AB}} b_i g_i \equiv_\beta c$$

Thus, there exists an integer t such that

$$\sum_{x_i \in V_{AB}} b_i g_i + t\beta = c \tag{6}$$

The equation $R_2BX = R_2B'$ is implied by $BX = B'$. Thus, the satisfying assignment $x_i = g_i$ for all $x_i \in V_{AB} \cup V_{B\setminus A}$ satisfies $R_2BX = R_2B'$. By plugging in the values g_i for x_i in Equation 3 we get:

$$\sum_{x_i \in V_{AB}} e_i g_i + \sum_{x_i \in V_{B\setminus A}} f_i g_i = d \tag{7}$$

We can sum the equations 6, 7 to get

$$t\beta + \sum_{x_i \in V_{AB}} (b_i + e_i) g_i + \sum_{x_i \in V_{B\setminus A}} f_i g_i = c + d \tag{8}$$

We know that t, β are integers, g_i are integers for all $x_i \in V_{AB} \cup V_{B\setminus A}$, and from Lemma 8 it follows that $b_i + e_i$ is integer for $x_i \in V_{AB}$ and f_i is integer for $x_i \in V_{B\setminus A}$. It follows that the left hand side of Equation 8 is an integer. While the right hand side of Equation 8 is not an integer by Lemma 8. Thus, the above equation is the required contradiction. It follows that $I_\beta \wedge BX = B'$ are unsatisfiable.

3. By the definition of I_β it follows that I_β only contains common variables of $AX = A'$ and $BX = B'$. $\square$

A.3 Algorithm for Computing Interpolants for LDEs

Algorithm 1 Interpolation for Linear Diophantine Equations

Require: Systems of LDEs $AX = A'$ and $BX = B'$, $AX = A' \wedge BX = B'$ is unsatisfiable.
Ensure: Return an interpolant for $(AX = A', BX = B')$

1: $[R_1, R_2] \Leftarrow$ proof of unsatisfiability of $AX = A' \wedge BX = B'$
$\{R_1A + R_2B$ is integral and $R_1A' + R_2B'$ is not an integer$\}$
2: $PI \Leftarrow R_1AX = R_1A'$ $\{PI$ represents partial interpolant$\}$
3: PI can be written as

$$\sum_{x_i \in V_{A \setminus B}} a_i x_i + \sum_{x_i \in V_{AB}} b_i x_i = c$$

$\{V_{AB} \subseteq X$ represents variables occuring in both $AX = A', BX = B'$, while $V_{A \setminus B} \subseteq X$ represents variables occurring in only $AX = A'\}$
4: **if** $a_i = 0$ for all $x_i \in V_{A \setminus B}$ **then**
5: return PI {Interpolant is a LDE}
6: **else**
7: $\alpha \Leftarrow gcd\{a_i | x_i \in V_{A \setminus B}\}$ $\{\alpha$ is an integer$\}$
8: Let β be any integer that divides α. Let linear modular equation

$$I_\beta := \sum_{i \in V_{AB}} b_i x_i \equiv_\beta c$$

9: return I_β {Interpolant is a LME}
10: **end if**

B Proofs from Section 4

B.1 Proof of Theorem 3

In order to prove theorem 3 we reduce the given system of LMEs to an equisatisfiable system of LDEs. We then use theorem 1 about the satisfiability of LDEs in order to complete the proof.

Reduction of a system of LMEs to an equisatisfiable system of LDEs

Suppose we are given a system $CX \equiv_l D$ of linear modular equations:

$$\underbrace{\begin{bmatrix} c_{11} & \dots & c_{1n} \\ c_{21} & \dots & c_{2n} \\ \dots & & \\ c_{m1} & \dots & c_{mn} \end{bmatrix}}_{C} \underbrace{\begin{bmatrix} x_1 \\ . \\ . \\ x_n \end{bmatrix}}_{X} \equiv_l \underbrace{\begin{bmatrix} d_1 \\ d_2 \\ . \\ d_m \end{bmatrix}}_{D}$$

For each equation $\sum_j c_{ij}x_j \equiv_l d_i$ in $CX \equiv_l D$ we introduce a new **integer** variable v_i, to obtain a new equation (without modulo), given as follows:

$$\sum_{j=1}^{n} c_{ij}x_j + lv_i = d_i$$

The above equation is equi-satisfiable to the linear modular equation $\sum_j c_{ij}x_j \equiv_l d_i$. Let V denote the vector of variables $v_1, \ldots, v_m$. We call the new system of linear equations as $C'Z = D$, where Z denotes the concatenation of variable vectors X and V. Note that $C'Z = D$ is a system of linear diophantine equations.

$$\underbrace{\begin{bmatrix} c_{11} & \ldots & c_{1n} & l & 0 & \ldots & 0 \\ c_{21} & \ldots & c_{2n} & 0 & l & \ldots & 0 \\ \ldots & & & & & & \\ c_{m1} & \ldots & c_{mn} & 0 & 0 & \ldots & l \end{bmatrix}}_{C'} \quad \underbrace{\begin{bmatrix} x_1 \\ . \\ x_n \\ v_1 \\ . \\ v_m \end{bmatrix}}_{Z} = \underbrace{\begin{bmatrix} d_1 \\ . \\ . \\ d_m \end{bmatrix}}_{D}$$

Lemma 9 *The following are equivalent:*
(a) the system of linear modular equations $CX \equiv_l D$ has an integral solution
(b) the system of linear diophantine equations $C'Z = D$ has an integral solution.

Proof. The proof of the above lemma is elementary.

Theorem 3. *Let C be a rational matrix, D be a rational column vector, and l be a rational number. The system $CX \equiv_l D$ has no integral solution X if and only if there exists a rational row vector R such that RC is integral, lR is integral, and RD is not an integer.*

From lemma 9 and theorem 1 the following are equivalent:

(a) linear modular equations $CX \equiv_l D$ has no integral solution
(b) linear diophantine equations $C'Z = D$ has no integral solution
(c) There exists a row vector R such that RC' is integral and RD is not an integer.

We show that the property of R in (c) is equivalent to "(d) RC is integral, lR is integral, and RD is not an integer".

Let $R = [r_1, \ldots, r_m]$ then

$$RC' = \left[\sum_{i=1}^{m} r_i c_{i1}, \sum_{i=1}^{m} r_i c_{i2}, \ldots, \sum_{i=1}^{m} r_i c_{in}, \;\; lr_1, \ldots, lr_i, \ldots, lr_m\right]$$

$$RC' = [RC, lR]$$

Thus, RC' is integral if and only if RC and lR are integral. This shows (c) is equivalent to (d). Thus, (a) is equivalent to (d) as required by the proof. □

B.2 Proof of Theorem 4

Recall that $V_{A\setminus B}$ denotes the set of variables that occur **only** in $AX \equiv_l A'$ (and not in $BX \equiv_l B'$) and V_{AB} denotes the set of variables that occur in both $AX \equiv_l A'$ and $BX \equiv_l B'$. The rational row vector $R = [R_1, R_2]$ is a proof of unsatisfiability of $AX \equiv_l A' \wedge BX \equiv_l B'$ such that

$$R_1A + R_2B \qquad \text{is integral} \tag{9}$$
$$lR = [lR_1, lR_2] \qquad \text{is integral} \tag{10}$$
$$R_1A' + R_2B' \qquad \text{is not an integer.} \tag{11}$$

Lemma 10 *The coefficient of $x_i \in V_{A\setminus B}$ in R_1AX is an integer.*

Proof. By definition of $V_{A\setminus B}$ the coefficient of $x_i \in V_{A\setminus B}$ is zero in R_2BX. Thus, the coefficient of $x_i \in V_{A\setminus B}$ is the same in R_1AX and $(R_1A + R_2B)X$. We know $R_1A + R_2B$ is integral from equation 9. So the coefficient of $x_i \in V_{A\setminus B}$ in R_1AX is an integer. □

Theorem 4. *We assume $l \neq 0$. Let S_1 denote the set of non-zero coefficients of $x_i \in V_{A\setminus B}$ in R_1AX. Let S_2 denote the set of all non-zero elements of row vector lR_1. If $S_2 = \emptyset$, then the interpolant for $(AX \equiv_l A', BX \equiv_l B')$ is a trivial LME $0 \equiv_l 0$. Otherwise, let $S_2 \neq \emptyset$. Let α denote the gcd of numbers in $S_1 \cup S_2$. (a) α is an integer and $\alpha > 0$. (b) Let β be any integer that divides α. Let $U = \frac{l}{\beta}R_1$. Then $UAX \equiv_l UA'$ is an interpolant for $(AX \equiv_l A', BX \equiv_l B')$.*

Proof. **$S_2 = \emptyset$:** If $S_2 = \emptyset$ it follows that all elements of lR_1 are zero. Since $l \neq 0$, R_1 must be a zero vector. It follows that R_1A is a zero vector and $R_1A' = 0$. Using equation 9 and R_1A is a zero vector, it follows that R_2B is integral. Using equation 11 and $R_1A' = 0$, it follows that R_2B' is not an integer. Thus, $BX \equiv_l B'$ is itself unsatisfiable with R_2 as the proof of unsatisfiability. In this case we can simply take `true` as the interpolant for the pair $(AX \equiv_l A', BX \equiv_l B')$. The interpolant `true` can be expressed as a trivial LME $0 \equiv_l 0$.

$S_2 \neq \emptyset$: We first show that α is an integer. Since lR_1 is integral (see equation 10) all elements of S_2 are non-zero integers. All elements of S_1 are non-zero integers due to Lemma 10. Thus, $S_1 \cup S_2$ is a set of non-zero integers. Since $S_2 \neq \emptyset$ there exists at least one element in $S_1 \cup S_2$. α is the gcd of the numbers in $S_1 \cup S_2$. So α is a non-zero integer and by definition of gcd α is positive.

Let β be any integer that divides α. Note that $\beta \neq 0$ as $\alpha \neq 0$. We define

$$I_\beta := UAX \equiv_l UA' \qquad \text{where} \qquad U = \frac{l}{\beta}R_1. \tag{12}$$

We need to show that I_β is an interpolant for the pair $(AX \equiv_l A', BX \equiv_l B')$.

(a) To show $AX \equiv_l A' \Rightarrow I_\beta$. If we show that U is integral, then by lemma 4 it follows that $AX \equiv_l A' \Rightarrow UAX \equiv_l UA'$ and thus $AX \equiv_l A' \Rightarrow I_\beta$. We need to show that U is integral.

Recall from equation 10 that lR_1 is integral. By definition of α it follows that α divides every element in S_2 or the row vector lR_1. Since β divides α, β divides every element in lR_1. So $\frac{lR_1}{\beta} = \frac{l}{\beta}R_1 = U$ is an integral vector.

(b) To show $I_\beta \wedge (BX \equiv_l B')$ is unsatisfiable. Observe that $I_\beta \wedge (BX \equiv_l B')$ is another system of LMEs

$$\begin{bmatrix} UA \\ B \end{bmatrix} X \equiv_l \begin{bmatrix} UA' \\ B' \end{bmatrix}$$

We show that the row vector $[\frac{\beta}{l}, R_2]$ serves as the proof of unsatisfiability of $I_\beta \wedge (BX \equiv_l B')$. We will check the conditions in the definition of proof of unsatisfiability.
- To show

$$[\frac{\beta}{l}, R_2] \begin{bmatrix} UA \\ B \end{bmatrix} \qquad \text{is integral}$$

The above product is equal to $\frac{\beta}{l}(UA) + R_2B = R_1A + R_2B$. By equation 9 we know that $R_1A + R_2B$ is integral.

- To show that $l[\frac{\beta}{l}, R_2] = [\beta, lR_2]$ is integral. From equation 10, lR_2 is integral and β is an integer by definition.

- To show

$$[\frac{\beta}{l}, R_2] \begin{bmatrix} UA' \\ B' \end{bmatrix} \qquad \text{is not an integer}$$

The above product is equal to $\frac{\beta}{l}(UA') + R_2B' = R_1A' + R_2B'$. By equation 11 we know that $R_1A' + R_2B'$ is not an integer.

We conclude that $[\frac{\beta}{l}, R_2]$ is a proof of unsatisfiability of $I_\beta \wedge (BX \equiv_l B')$. Thus, $I_\beta \wedge (BX \equiv_l B')$ is unsatisfiable.

(c) To show that I_β only contains variables that are common to both $(AX \equiv_l A', BX \equiv_l B')$. Since I_β is obtained by a linear combination of equations from $AX \equiv_l A'$, we can write I_β as follows:

$$\underbrace{\sum_{x_i \in V_{A \setminus B}} a_i x_i + \sum_{x_i \in V_{AB}} b_i x_i}_{UAX} \equiv_l \underbrace{c}_{UA'} \tag{13}$$

where all coefficients a_i, b_i and $c = UA'$ are rational numbers.

We will show that the coefficient a_i of each $x_i \in V_{A \setminus B}$ in equation 13 is divisible by l. This will in turn show that

$$\sum_{x_i \in V_{A \setminus B}} a_i x_i \equiv_l 0 \tag{14}$$

since x_i are integer variables. This will allow I_β to be written in an equivalent manner (containing only variables from V_{AB}) as follows:

$$\sum_{x_i \in V_{AB}} b_i x_i \equiv_l c.$$

We now show that the coefficient a_i of each $x_i \in V_{A \setminus B}$ in equation 13 is divisible by l. Recall, that

$$I_\beta := UAX \equiv_l UA' \qquad \text{where} \qquad U = \frac{l}{\beta} R_1 \text{ and } \beta \text{ divides } \alpha. \tag{15}$$

By definition α divides every element in S_1
$\Rightarrow \alpha$ divides the coefficient of each $x_i \in V_{A\setminus B}$ in R_1AX
$\Rightarrow \beta$ divides the coefficient of each $x_i \in V_{A\setminus B}$ in R_1AX.
$\Rightarrow$ the coefficient of $x_i \in V_{A\setminus B}$ in $\frac{1}{\beta}R_1AX$ is an integer.
$\Rightarrow$ the coefficient of $x_i \in V_{A\setminus B}$ in $l \times \frac{1}{\beta}R_1AX$ is divisible by l.
$\Rightarrow$ the coefficient of $x_i \in V_{A\setminus B}$ in UAX is divisible by l (as $U = \frac{l}{\beta}R_1$)
The coefficient of $x_i \in V_{A\setminus B}$ in UAX is simply a_i (equation 13). So l divides a_i. $\square$

Degenerate case $l = 0$. Let $AX \equiv_l A'$ be a system of LMEs. For $l = 0$, $AX \equiv_l A'$ is equivalent to a system of LDEs $AX = A'$. In order to see this, consider an LME $\sum_{i=1}^{n} a_i x_i \equiv_0 b$. This LME is satisfied if and only if $\sum_{i=1}^{n} a_i x_i - b = 0 \times \lambda$, for some integer λ. Thus, the LME $\sum_{i=1}^{n} a_i x_i \equiv_0 b$ is equivalent to the LDE $\sum_{i=1}^{n} a_i x_i = b$.

Suppose $AX \equiv_0 A' \wedge BX \equiv_0 B'$ is unsatisfiable. Then the interpolant for $(AX \equiv_0 A', BX \equiv_0 B')$ can be obtained by computing the interpolant for the pair of LDEs $(AX = A', BX = B')$.

C Proof of Corollary 1

Corollary 1. *Given $CX = D$ where C, D are rational matrices, and C has full row rank. Let $[E\ \ 0]$ denote the Hermite normal form (HNF) of C. If $CX = D$ has no integral solution, then $E^{-1}D$ is not integral (due to lemma 5). Suppose the i^{th} entry in $E^{-1}D$ is not an integer. Let R' denote the i^{th} row in E^{-1}. Then*
(a) $R'D$ is not an integer
(b) $R'C$ is integral
Thus, R' serves as the required proof of unsatisfiability of $CX = D$.

Proof. (a) Follows from the definition of R'
(b) We know that

$$CU = [E\ \ 0]$$

where U is a unimodular matrix. Since E is invertible (by definition of HNF) we can multiply both sides of the above equation by E^{-1} to obtain

$$E^{-1}CU = E^{-1}[E\ \ 0].$$

The above equation simplifies to

$$E^{-1}CU = [I\ \ 0]$$

where I is the identity matrix. Since U is unimodular its inverse (U^{-1}) exists and it is a unimodular matrix. Multiply both sides of the above equation by U^{-1} to obtain

$$E^{-1}CUU^{-1} = [I\ \ 0]U^{-1}.$$

The above equation simplifies to

$$E^{-1}C = [I\ \ 0]U^{-1}.$$

Since U^{-1} is unimodular the right hand side of the above equation has integral entries. Thus, the left hand side $E^{-1}C$ is integral. In particular the ith row in $E^{-1}C$ is integral. Observe that the ith row in $E^{-1}C$ is simply $R'C$. Thus, $R'C$ is integral. $\square$

D Proof of Lemma 6

We need to introduce cutting-plane proof system [29, 7] in order to prove this lemma. Suppose we are given a system of integer linear inequalities $AX \leq B$, where A, B are rational matrices and X is a column vector of integer variables. The following inference rules allow us to derive new inequalities that are implied by $AX \leq B$.

`nonneg_lin_comb`: We can take a non-negative linear combination of inequalities to derive a new inequality.

$$\frac{AX \leq B}{RAX \leq RB} \qquad R \geq 0$$

(R is a rational row vector whose each element is non-negative.)

`rounding`: If we have a linear inequality $EX \leq F$ such that all coefficients in E are integers ($E \in \mathbb{Z}^n$), then we can round down the right hand side F.

$$\frac{EX \leq F}{EX \leq \lfloor F \rfloor} \qquad E \in \mathbb{Z}^n$$

($EX \leq F$ in the above rule represents a single inequality and not a system of inequalities. E is a row vector containing n integers.) We say an application of the `rounding` rule is *redundant* if $F = \lfloor F \rfloor$ in the above inference rule.

`weak_rhs`: Given $F \leq F'$ and a linear inequality $EX \leq F$ we can derive $EX \leq F'$

$$\frac{EX \leq F}{EX \leq F'} \qquad F \leq F'$$

We say an application of the `weak_rhs` rule is *redundant* if $F = F'$ in the above inference rule.

A *cutting plane proof* of an inequality $EX \leq F$ from $AX \leq B$ is a sequence of inequalities $E_1X \leq F_1, \ldots, E_lX \leq F_l$ such that

$$\frac{AX \leq B, E_1X \leq F_1, \ldots, E_{i-1}X \leq F_{i-1}}{E_iX \leq F_i} \qquad \texttt{nonneg\_lin\_comb or rounding}$$

for each $i = 1, \ldots, l$ and each step is an application of the `nonneg_lin_comb` or the `rounding` inference rules ($E_1, \ldots, E_l$ are rational row vectors and $F_1, \ldots, F_l$ are rational numbers). We do not need the `weak_rhs` rule anywhere, except possibly as the last step in a cutting plane proof.

$$\frac{E_lX \leq F_l}{EX \leq F} \qquad E = E_l, F_l \leq F'.$$

The cutting plane proof system provides a sound and complete inference system for integer linear inequalities. This is stated formally in the following theorem.

Theorem 6 *(Schrijver [29]) We are given a system of integer linear inequalities $AX \leq B$, where A, B are rational matrices and X is a column vector of integer variables. Let $EX \leq F$ be an inequality, where E is a rational row vector and F is a rational number.*

1. $AX \leq B$ has an integral solution and $AX \leq B$ implies $EX \leq F$ if and only if there is a cutting plane proof of $EX \leq F$ from $AX \leq B$.
2. $AX \leq B$ has no integral solution if and only if then there is a cutting plane proof of $0 \leq -1$ from $AX \leq B$.

We need to prove the following:
Lemma 6: *The following are equivalent:*
1. A system of LDEs $AX = B$ implies a LDE $EX = F$
2. $AX = B$ has no integral solution or there exists a rational row vector R such that $E = RA$ and $F = RB$.

Proof. (2) $\Rightarrow$ (1) is straightforward.
(1) $\Rightarrow$ (2): Given $AX = B$ implies a linear equation $EX = F$. If $AX = B$ has no integral solution we are done, that is, (2) holds. Otherwise, assume that $AX = B$ has an integral solution.

We can write $AX = B$ as an equivalent system of inequalities $AX \leq B \wedge -AX \leq -B$. The cutting plane (CP) proof rules provide a complete inference system for integer linear inequalities. We can write the LDE $EX = F$ as $EX \leq F \wedge -EX \leq -F$. The system of linear inequalities $AX \leq B \wedge -AX \leq -B$ implies $EX \leq F \wedge -EX \leq -F$. Let us consider the CP proof of $EX \leq F$ from the inequalities $AX \leq B \wedge -AX \leq -B$. We show that the inference rules used in this proof will only involve `nonneg_linear_comb` rule. Any application of `rounding` or `weak_rhs` rule will either be redundant or will lead to a contradiction. The later case is not possible because $AX = B$ or the equivalent system of inequalities has an integral solution.

Consider the first application of `rounding` in the CP proof of $EX \leq F$.

$$\frac{E_iX \leq F_i}{E_iX \leq \lfloor F_i \rfloor} \qquad E_i \in \mathbb{Z}^n$$

Since all the rules used to derive $E_iX \leq F_i$ are non negative linear combination rules, we can combine all steps used to derive $E_iX \leq F_i$ by a single application of the `nonneg_lin_comb` rule. That is, we can find rational row vector $[R_1, R_2]$ such that

$$\frac{\begin{bmatrix} A \\ -A \end{bmatrix} X \leq \begin{bmatrix} B \\ -B \end{bmatrix}}{\underbrace{[R_1, R_2]\begin{bmatrix} A \\ -A \end{bmatrix} X}_{E_iX} \leq \underbrace{[R_1, R_2]\begin{bmatrix} B \\ -B \end{bmatrix}}_{F_i}} \qquad [R_1, R_2] \geq 0$$

where R_1, R_2 are non-negative, $E_i = R_1A + R_2(-A)$ and $F_i = R_1B + R_2(-B)$. We can also derive $-E_iX \leq -F_i$ by taking a non negative linear combination of $AX \leq B \wedge -AX \leq -B$ using $[R_2, R_1]$. If $F_i = \lfloor F_i \rfloor$ then the application of rounding rule

$$\frac{E_iX \leq F_i}{E_iX \leq \lfloor F_i \rfloor} \qquad E_i \in \mathbb{Z}^n$$

is redundant. Otherwise, let $\lfloor F_i \rfloor = k (\neq F_i)$ and

$$\frac{E_iX \leq F_i}{E_iX \leq k}$$

Since $\lfloor -F_i \rfloor = -k-1$. We apply apply rounding to $-E_iX \leq -F_i$ to obtain

$$\frac{-E_iX \leq -F_i}{-E_iX \leq -k-1} \qquad -E_i \in \mathbb{Z}^n$$

By combining the above two equations ($E_iX \leq k$ and $-E_iX \leq -k-1$) we obtain an equation $0 \leq -1$. But this means that the original system of inequalities $AX \leq B \wedge -AX \leq -B$ has no integral solution, which contradicts our assumption. Thus, the first application of the `rounding` rule in the CP proof must be redundant. Using similar reasoning (induction on the length of the proof) we can conclude that all applications of `rounding` in the CP proof must be redundant.

In the CP proof system described above there can be only one application of `weak_rhs` rule as the last step in a CP proof. We now show that the application of `weak_rhs` at the end of the CP proof must be redundant.

$$\frac{EX \leq F_l}{EX \leq F} \qquad F_l \leq F.$$

If $F_l = F$, then the application of `weak_rhs` is redundant. Otherwise, suppose $F_l < F$. Recall, that $-EX \leq -F$ is also an implied inequality of the original system. We can add $-EX \leq -F$ and $EX \leq F_l$ to obtain $0 \leq F_l - F$. Since $F_l < F$ we can divide $0 \leq F_l - F$ by positive rational number $F - F_l$, to obtain the equation $0 \leq -1$. But this is a contradiction.

Thus, the cutting plane proof of $EX \leq F$ can only involve redundant applications of `rounding` or `weak_rhs` rules. These applications of `rounding` or `weak_rhs` rules can be removed to obtain a derivation of $EX \leq F$ that only involves `nonneg_linear_comb` rule. All applications of `nonneg_linear_comb` rule in a CP proof can be combined to obtain a vector $[S_1, S_2]$ such that

$$\frac{\begin{bmatrix} A \\ -A \end{bmatrix} X \leq \begin{bmatrix} B \\ -B \end{bmatrix}}{\underbrace{[S_1,S_2]\begin{bmatrix} A \\ -A \end{bmatrix}}_{EX} X \leq \underbrace{[S_1,S_2]\begin{bmatrix} B \\ -B \end{bmatrix}}_{F}} \qquad [S_1, S_2] \geq 0$$

where S_1, S_2 are non-negative, $E = S_1A + S_2(-A)$ and $F = S_1B + S_2(-B)$. (Note that a proof of $-EX \leq -F$ can be obtained by taking a non negative linear combination of $AX \leq B, -AX \leq -B$ using $[S_2, S_1]$.) Thus, there exists a rational vector $R = S_1 - S_2$ such that $E = RA$ and $F = RB$. This shows (2) holds. $\square$

E Proof of Lemma 7

We use the following result in the proof.

Theorem 7 *(Schrijver [29]) Let $AX = B$ be a system of LDEs, where A, B are rational matrices and X is a column vector of n integer variables. If $AX = B$ is satisfiable (has an integral solution), then we can find in polynomial time integral vectors $X_0, \ldots, X_t \in \mathbb{Z}^n$ such that*

$$\{X | AX = B; X \text{ integral}\} = \{X_0 + \lambda_1 X_1 + \ldots + \lambda_t X_t | \lambda_1, \ldots, \lambda_t \in \mathbb{Z}\}$$

with $X_1, \ldots, X_t$ linearly independent. (We think of $X_0, X_1, \ldots, X_t \in \mathbb{Z}^n$ as column vectors.)

Example 14 Consider a system of LDEs $AX = B$:

$$\begin{bmatrix} 2 & 6 & 3 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$$

The set S of solutions to $AX = B$ is given as:

$$S = \left\{ \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix} + \lambda_1 \begin{bmatrix} -3 \\ 3 \\ -4 \end{bmatrix} \middle| \lambda_1 \in \mathbb{Z} \right\} = \left\{ \begin{bmatrix} 2 - 3\lambda_1 \\ 3\lambda_1 \\ -4\lambda_1 \end{bmatrix} \middle| \lambda_1 \in \mathbb{Z} \right\}$$

Lemma 7: *Let $AX = B$ denote a system of LDEs, where A, B are rational matrices and X is a column vector of integer variables. Let $C_iX = D_i$ denote a LDE for $1 \leq i \leq m$ (C_i is a rational row vector and D_i is a rational number). The following are equivalent:*
1. $AX = B$ implies $\bigvee_{i=1}^{m} C_iX = D_i$
2. There exists a $1 \leq k \leq m$ such that $AX = B$ implies $C_kX = D_k$.

Proof. (2) $\Rightarrow$ (1): This direction of the proof is straightforward.
(1) $\Rightarrow$ (2): If $AX = B$ has no integral solution, then $AX = B$ implies any linear equation. Thus, (2) holds.

Assume that $AX = B$ has an integral solution. In this case we can use the theorem 7 and write the set S of all integral solutions to $AX = B$ as

$$S := \{X_0 + \lambda_1 X_1 + \ldots + \lambda_t X_t | \lambda_1, \ldots, \lambda_t \in \mathbb{Z}\}$$

where $X_0, X_1, \ldots, X_t \in \mathbb{Z}^n$ (assuming X has size $n \times 1$).

By substituting $X = X_0 + \lambda_1 X_1 + \ldots + \lambda_t X_t$ (with $\lambda_1, \ldots, \lambda_t$ as symbolic variables) in $C_iX - D_i$ we obtain

$$C_i(X_0 + \lambda_1 X_1 + \ldots + \lambda_t X_t) - D_i.$$

Since $C_iX_0, \ldots, C_iX_t$ are scalars (rational numbers), the difference $C_iX - D_i$ for $X \in S$ is a linear expression in $\lambda_1, \ldots, \lambda_t$. We denote the difference $C_iX - D_i$ for $X \in S$ by δ_i. It follows that

$$\left.\begin{array}{rcl} \delta_1 & = & u_{10} + u_{11}\lambda_1 + \ldots + u_{1t}\lambda_t \\ & \ldots & \\ \delta_i & = & u_{i0} + u_{i1}\lambda_1 + \ldots + u_{it}\lambda_t \\ & \ldots & \\ \delta_m & = & u_{m0} + u_{m1}\lambda_1 + \ldots + u_{mt}\lambda_t \end{array}\right\} EQ$$

where u_{ij} are rational numbers, $\lambda_1, \ldots, \lambda_t$, $\delta_1, \ldots, \delta_m$ are symbolic variables. An integral assignment $\lambda_1 = \beta_1, \ldots, \lambda_t = \beta_t$ where $\beta_1, \ldots, \beta_t \in \mathbb{Z}$ gives a solution $X_\beta \in \mathbb{Z}^n$ to $AX = B$ ($X_\beta \in S$). If δ_i evaluates to zero for $\lambda_1 = \beta_1, \ldots, \lambda_t = \beta_t$, then X_β satisfies the LDE $C_iX = D_i$. Otherwise, X_β does not satisfy the LDE $C_iX = D_i$.

We consider two cases.
Case 1: If for some $1 \leq k \leq m$, $u_{k0} = \ldots = u_{kt} = 0$, then $\delta_k = 0$. That is, every $X \in S$ satisfies $C_kX = D_k$. Therefore, $AX = B$ implies $C_kX = D_k$. In this case (2) holds.

Case 2: For all $1 \leq k \leq m$ there is a $0 \leq j \leq t$ such that $u_{kj} \neq 0$. We show that case 2 cannot arise using proof by contradiction. We will give an algorithm for assigning integral values to $\lambda_1, \ldots, \lambda_t$ such that $\delta_1 \neq 0, \ldots, \delta_m \neq 0$. In other words, we will show that there exists an $X' \in S$ such that $C_i X' \neq D_i$ for all $1 \leq i \leq m$. This will mean that $AX = B$ does not imply $\vee_{i=1}^{m} C_i X = D_i$, leading to a contradiction.

It is convenient to think of expressions for $\delta_1, \ldots, \delta_m$ as a system of equations in $\delta_1, \ldots, \delta_m, \lambda_1, \ldots, \lambda_t$. We denote this system of equations as EQ.

We now give an algorithm for assigning integral values to $\lambda_1, \ldots, \lambda_t$ such that $\delta_1 \neq 0, \ldots, \delta_m \neq 0$. Our algorithm will assign λ_i before λ_{i+1} for each $1 \leq i \leq m-1$.

Let $EQ_0 \subseteq EQ$ denote the equations that do not contain any variables $\lambda_1, \ldots, \lambda_t$. If $\delta_k = u_{k0}$ is an equation in EQ_0, then we know that $u_{k0} \neq 0$ (by case 2 assumption). Thus, $C_k X \neq D_k$ for any $X \in S$. Alternatively, $AX = B$ cannot imply $C_k X = D_k$. We can safely ignore the equations in EQ_0 for the rest of the proof.

Let $EQ_i \subseteq EQ$ for $1 \leq i \leq t$ denote the set of equations which contain only variables $\lambda_1, \ldots, \lambda_i$ such that the coefficient of λ_i is not zero (coefficients of $\lambda_1, \ldots, \lambda_{i-1}$ can be zero).

We now describe an algorithm for assigning integer values to λ_i for $1 \leq i \leq t$. The algorithm uses EQ_i to assign a value to λ_i. Suppose we have assigned integral values $\alpha_1, \ldots, \alpha_{i-1}$ to $\lambda_1, \ldots, \lambda_{i-1}$, respectively. If $EQ_i = \emptyset$, then assign an arbitrary integer value α_i to λ_i. Otherwise, substitute $\lambda_1 = \alpha_1, \ldots, \lambda_{i-1} = \alpha_{i-1}$ in EQ_i to obtain a system of equations EQ_i'. A representative equation in EQ_i' is

$$\delta_l = v_{l0} + u_{li}\lambda_i \qquad\qquad u_{li} \neq 0$$

where v_{l0} is a rational number and u_{li} is a non-zero rational number by definition of EQ_i. We want to assign λ_i such that $\delta_l \neq 0$ for every equation $\delta_l = v_{l0} + u_{li}\lambda_i$ in EQ_i'. This can be done by assigning λ_i any integer value that is different from $\frac{-v_{l0}}{u_{li}}$. Let

$$\lambda_i := \alpha_i \qquad \text{where} \qquad \alpha_i \in \mathbb{Z} \qquad \text{and} \qquad \alpha_i \notin \left\{ \frac{-v_{l0}}{u_{li}} | l \in EQ_i' \right\}$$

where $l \in EQ_i'$ is a short form of saying that equation $\delta_l = v_{l0} + u_{li}\lambda_i$ is in EQ_i'. We can always find a suitable α_i because the set of integers has infinite cardinality (and we have a finite set of rational numbers/integers that cannot be assigned to λ_i).

Let $\delta_l = u_{l0} + \sum_{j=1}^{i} u_{lj}\lambda_j$ denote an equation in $EQ_1 \cup \ldots \cup EQ_i$. The following invariant holds after λ_i is assigned α_i: if $\lambda_1 = \alpha_1, \ldots, \lambda_i = \alpha_i$ is substituted in $\delta_l = u_{l0} + \sum_{j=1}^{i} u_{lj}\lambda_j$, then $\delta_l \neq 0$.

Thus, once we have assigned $\lambda_1 = \alpha_1, \ldots, \lambda_t = \alpha_t$ using the above algorithm we have $\delta_1 \neq 0, \ldots, \delta_m \neq 0$. Let $X' \in S$ be an integral solution to $AX = B$ given by $\lambda_1 = \alpha_1, \ldots, \lambda_t = \alpha_t$. Then $\delta_i = C_i X' - D_i \neq 0$ for each $1 \leq i \leq m$. That is, $AX = B$ does not imply $\vee_{i=1}^{m} C_i X = D_i$, leading to a contradiction. Thus, Case 2 cannot arise. □

F Proof of Theorem 5

In addition to lemmas 6,7 we will use the following theorem.

Theorem 8 *(Schrijver [29]) Let A be a rational matrix, B be a rational column vector, C be a rational row vector. Assume that the system $AX = B$ has a rational solution. Then $AX = B$ implies (over rationals) $CX = D$ if and only if there is a row vector R such that $RA = C$ and $RB = D$.*

Theorem 5. *Let F denote $AX = B \wedge \bigwedge_{i=1}^{m} C_i X \neq D_i$. The following are equivalent:*
1. F has no integral solution
2. F has no rational solution or $AX = B$ has no integral solution.

Proof. (2) $\Rightarrow$ (1) is straightforward.
(1) $\Rightarrow$ (2): Given F has no integral solution. If $AX = B$ has no integral solution, then (2) holds. Otherwise, assume $AX = B$ has an integral solution. Since F has no integral solution, every integral solution to $AX = B$ must satisfy $C_i X = D_i$ for some $1 \leq i \leq m$. That is,

$$AX = B \Rightarrow \bigvee_{i=1}^{m} C_i X = D_i$$

By lemma 7 it follows that there exists a $1 \leq k \leq m$ such that

$$AX = B \Rightarrow C_k X = D_k$$

By lemma 6 (and our assumption that $AX = B$ has an integral solution) it follows that there exists a rational row vector R such that

$$C_k = RA \qquad \text{and} \qquad D_k = RB$$

Using the vector R and theorem 8 we can conclude that $AX = B$ implies $C_k X = D_k$ over rationals. So

$$AX = B \wedge C_k X \neq D_k$$

is unsatisfiable over rationals, or

$$AX = B \wedge \bigwedge_{i=1}^{m} C_i X \neq D_i$$

is unsatisfiable over rationals. Thus, F is unsatisfiable over rationals and (2) holds. □

G Interpolants for Linear Diophantine Equations and Disequations (LDEs+LDDs)

We use the following theorem.

Theorem 9 *(Schrijver [29]) Let A be a rational matrix, B be a rational column vector. The system $AX = B$ has no rational solution if and only if there exists a rational row vector R such that $RA = 0$ and $RB \neq 0$.*

Let $F \wedge G$ be systems of LDEs+LDDs.

$$F \;:=\; AX = B \wedge \bigwedge_{i} C_i X \neq D_i$$

$$G \;:=\; A'X = B' \wedge \bigwedge_{j} C'_j X \neq D'_j$$

$F \wedge G$ represents another system of LDEs+LDDs. Suppose $F \wedge G$ is unsatisfiable (no integral solution). In this case we want to compute an interpolant for the pair (F, G). We divided this problem into two cases in Section 6. We describe Case 1 below.

By case 1 assumption we know that $F \wedge G$ has no rational solution. We want to compute an interpolant for (F, G). The interpolant for (F, G) can be obtained by using the techniques discussed in [24, 33, 28, 10]. For completeness we show how to obtain an interpolant for (F, G) by considering three sub-cases.

Case 1.1: $AX = B \wedge A'X = B'$ has no rational solution. Using theorem 9 there exists a row vector $[R_1, R_2]$ such that

$$R_1A + R_2A' = 0$$
$$R_1B + R_2B' \neq 0$$

In this case an interpolant for the pair (F, G) is the linear equation $R_1AX = R_1B$. One can verify that $R_1AX = R_1B$ satisfies all the conditions required by the definition of interpolants.

We describe Case 1.2 and Case 1.3 next. Since $F \wedge G$ is unsatisfiable over rationals we have

$$AX = B \wedge A'X = B' \Rightarrow (\bigvee_i C_iX = D_i \vee \bigvee_j C'_jX = D'_j) \tag{16}$$

The above implication holds for any rational X. We know that if a set of rational linear arithmetic constraints Γ imply a disjunction of linear equations $\bigvee_{i=1}^{m} Eq_i$, then for some $1 \leq k \leq m$, Γ implies Eq_k. This is due to *convexity* of rational linear arithmetic [25].

Due to convexity $AX = B \wedge A'X = B'$ will imply either an equality belonging to $\bigvee_i C_iX = D_i$ or an equality belonging to $\bigvee_j C'_jX = D'_j$ in equation 16. This gives Case 1.2 and Case 1.3.

Case 1.2: For some j, $AX = B \wedge A'X = B' \Rightarrow C'_jX = D'_j$.
Using theorem 8 there exists a row vector $[R_1, R_2]$ such that

$$\begin{aligned} R_1A + R_2A' &= C'_j \\ R_1B + R_2B' &= D'_j. \end{aligned}$$

In this case an interpolant for (F, G) is the linear equation $R_1AX = R_1B$. One can verify that $R_1AX = R_1B$ satisfies all the conditions required by the definition of interpolants.

Case 1.3: For some i, $AX = B \wedge A'X = B' \Rightarrow C_iX = D_i$.
In the above two cases (1.1 and 1.2) the interpolant is a linear equation. In this case the interpolant will be a linear disequation. Using theorem 8 there exists a row vector $[R_1, R_2]$ such that

$$R_1A + R_2A' = C_i$$
$$R_1B + R_2B' = D_i$$

Let V_{FG} denote the variables that occur in both F and G and let $V_{F\setminus G}$ denote the variables that occur only in F (and not in G).

Observe that $R_1AX = R_1B$ can be written as follows:

$$\sum_{x_i \in V_{F\setminus G}} a_ix_i + \sum_{x_i \in V_{FG}} b_ix_i = k$$

Similarly, $C_iX = D_i$ can be written as follows:

$$\sum_{x_i \in V_{F\setminus G}} a_i x_i + \sum_{x_i \in V_{FG}} c_i x_i = D_i$$

Observe that the variables $x_i \in V_{F\setminus G}$ have same coefficients in R_1AX and C_iX. This is because $C_i = R_1A + R_2A'$ and the coefficients of $x_i \in V_{F\setminus G}$ in $R_2A'X$ is zero.

We can write $C_iX \neq D_i$ as

$$\sum_{x_i \in V_{F\setminus G}} a_i x_i + \sum_{x_i \in V_{FG}} c_i x_i \neq D_i$$

Note that F implies $R_1AX = R_1B$ and $C_iX \neq D_i$. Thus, F implies the disequation obtained by subtracting $R_1AX = R_1B$ and $C_iX \neq D_i$.

$$\sum_{x_i \in V_{FG}} b_i x_i - \sum_{x_i \in V_{FG}} c_i x_i \neq k - D_i$$

The above equation is the required interpolant. It it implied by F and only contains variables common to F, G. One can show that above disequation is $R_2A'X \neq R_2B'$. Since G implies $R_2A'X = R_2B'$ the above equation is unsatisfiable with G.

H Handling of Linear Modular Disequations

Lemma 11 *The problem of deciding whether a system (conjunction) of linear modular disequations (LMDs) have an integral solution is NP-hard.*

Proof. We reduce a well known NP-hard problem 3-SAT to a system of LMDs denoted by $\mathcal{L}$. Let the variables in 3-SAT problem be $z_1, \ldots, z_n$. For each variable z_i in the 3-SAT problem we introduce two integer variables x_i and x_i' in $\mathcal{L}$, where x_i represents the literal z_i and x_i' represents the literal $\bar{z_i}$.

The modulus of LMDs in $\mathcal{L}$ will be four. We first express the constraints that $x_i \equiv_4 1$ and $x_i' \equiv_4 0$ or $x_i \equiv_4 0$ and $x_i' \equiv_4 1$. This done by means of the following LMDs.

$$\begin{aligned}\mathcal{L}_1 \quad := \quad & \bigwedge_{i=1}^{n} \neg(x_i \equiv_4 x_i') \quad \wedge \quad \bigwedge_{i=1}^{n} \neg(x_i \equiv_4 2) \wedge \bigwedge_{i=1}^{n} \neg(x_i \equiv_4 3) \wedge \\ & \bigwedge_{i=1}^{n} \neg(x_i' \equiv_4 2) \wedge \bigwedge_{i=1}^{n} \neg(x_i' \equiv_4 3)\end{aligned}$$

Now consider any clause $u \vee v \vee w$ in the given 3-SAT formula, where $u, v, w \in \{z_1, \ldots, z_n, \bar{z_1}, \ldots, \bar{z_n}\}$. Let $\delta(u)$ map the literal u to the corresponding variable in $\mathcal{L}$. For each clause $u \vee v \vee w$ in the 3-SAT formula, we generate the following LMD

$$\neg(\delta(u) + \delta(v) + \delta(w) \equiv_4 0).$$

The LMD above is falsified only when $\delta(u), \delta(v), \delta(w)$ are assigned $0 \ (mod\ 4)$. For all other assignment of values $\delta(u), \delta(v), \delta(w)$ the LMD is satisfied (captures the semantics of the clause).

Let the set of clauses in the 3-SAT formula be C.

$$\mathcal{L}_2 \quad := \quad \bigwedge_{(u \vee v \vee w) \in C} \neg(\delta(u) + \delta(v) + \delta(w) \equiv_4 0)$$

Let $\mathcal{L} = \mathcal{L}_1 \wedge \mathcal{L}_2$. Observe that the 3-SAT formula is satisfiable if and only if $\mathcal{L}$ is satisfiable. The reduction from the given 3-SAT formula to $\mathcal{L}$ is polynomial time. This establishes the NP-hardness of checking the satisfiability of conjunctions of LMDs. $\square$

H.1 Proofs of unsatisfiability and interpolants for LMDs

We can reduce a system of LMDs or LMEs+LMDs to a conjunction of atomic formulas in integer linear arithmetic (both problems are NP-hard) and use the *cutting-plane proof system* to obtain a proof of unsatisfiability. Pudlak's [27] algorithm can be used for obtaining interpolants.

I Obtaining polynomially sized cutting-plane proofs for LDEs

Given an unsatisfiable system of LDEs $AX = B$, a proof of unsatisfiability is a rational row vector R such that RA is integral, while RB is not an integer. We know that R can be obtained in polynomial time.

We show that using R we can obtain a polynomially sized cutting plane proof of unsatisfiability of $AX = B$. The cutting plane proof system was described in Appendix D. It consists of three inference rules `nonneg_lin_comb`, `rounding` and `weak_rhs`.

We first write $R = S_1 - S_2$, where both S_1, S_2 are non-negative row vectors. For example, we can write $[\frac{1}{2}, -\frac{3}{4}] = [\frac{1}{2}, 0] - [0, \frac{3}{4}]$.

We write $AX = B$ as $AX \leq B \wedge -AX \leq -B$. The cutting plane proof of unsatisfiability consists of following steps.

$$\frac{AX \leq B}{S_1 AX \leq S_1 B} \quad S_1 \geq 0 \qquad \texttt{nonneg\_lin\_comb}$$

$$\frac{-AX \leq -B}{-S_2 AX \leq -S_2 B} \quad S_2 \geq 0 \qquad \texttt{nonneg\_lin\_comb}$$

$$\frac{S_1 AX \leq S_1 B \qquad -S_2 AX \leq -S_2 B}{[S_1 - S_2]AX \leq [S_1 - S_2]B} \qquad \texttt{nonneg\_lin\_comb}$$

Since $R = [S_1 - S_2]$ we can write the above step as

$$\frac{S_1 AX \leq S_1 B \qquad -S_2 AX \leq -S_2 B}{RAX \leq RB} \qquad \texttt{nonneg\_lin\_comb}$$

Multiplying $AX \leq B$ by S_2 and $-AX \leq -B$ by S_1 we can derive

$$\frac{S_2 AX \leq S_2 B \qquad -S_1 AX \leq -S_1 B}{-RAX \leq -RB} \qquad \texttt{nonneg\_lin\_comb}$$

By definition of R we know that RB is not an integer. Let $\lfloor RB \rfloor = k$. Then $\lfloor -RB \rfloor = -k - 1$. Since RA is integral we can apply rounding to $RAX \leq RB$ and $-RAX \leq -RB$.

$$\frac{RAX \leq RB}{RAX \leq k} \qquad \texttt{rounding}$$

$$\frac{-RAX \leq -RB}{RAX \leq -k-1} \qquad \texttt{rounding}$$

The contradiction is obtained by summing $RAX \leq k$ and $RAX \leq -k-1$.

$$\frac{RAX \leq RB \qquad -RAX \leq -RB}{0 \leq -1} \qquad \texttt{nonneg\_lin\_comb}$$

Since R is polynomially sized the cutting plane proof is also polynomially sized.

J Using SMT solvers for obtaining a proof of unsatisfiability for LDEs/LMEs

We can determine if a system of LDEs $CX = D$ is unsatisfiable and obtain a proof of unsatisfiability (if applicable) by using decision procedures for (mixed) integer linear arithmetic in a black-box fashion. For example, one can use modern SMT solvers such as Yices [4] to obtain proofs of unsatisfiability. The idea is to encode the existence of a rational row vector R such that RC is integral and RD is not an integer in form of a formula that can be checked using existing decision procedures. This is motivated by the idea proposed in [28] for real and rational linear arithmetic. We illustrate the technique by means of an example.

Example 15 Consider the system of LDEs $CX = D$:

$$\begin{bmatrix} 1 & -2 & 0 \\ 1 & 0 & -2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

We use two rational variables r_1, r_2 to denote the proof of unsatisfiability $R = [r_1, r_2]$. We use three integer variables v_1, v_2, v_3 to express the constraint that RC is integral. We introduce another integer variable v_4 to express the constraint that $RD = r_2$ is not an integer.

$$P := (v_1 = r_1 + r_2) \wedge (v_2 = -2r_1) \wedge (v_3 = -2r_2) \wedge (v_4 < r_2) \wedge (r_2 < v_4 + 1)$$

If the decision procedure for integer linear arithmetic determines that P is satisfiable, then we get a proof of unsatisfiability for $CX = D$ by looking at the assignments to r_1, r_2. If P is unsatisfiable, it means that the system $CX = D$ is satisfiable.

We formalize the idea below. Suppose the sizes of C, X, D in the system of LDEs $CX = D$ are $m \times n, n \times 1, m \times 1$, respectively. The formula P contains:
- m rational variables $r_1, \ldots, r_m$ such that $R = [r_1, \ldots, r_m]$
- n integer variables $v_1, \ldots, v_n$ to express that each element of RC is integral.
- One integer variable v_{n+1} to express the constraint RD is not an integer by using two strict inequalities

Let $(RC)_i$ denote the ith element in the row vector RC. Then we have

$$P := \bigwedge_{i=1}^{n} v_i = (RC)_i \ \wedge \ (v_{n+1} < RD) \ \wedge \ (RD < v_{n+1} + 1)$$

The formula P is given to a SMT solver. If P is satisfiable, we get the required proof of unsatisfiability R. Otherwise, we know that the given system of LDEs is satisfiable.

The proof of unsatisfiability for a system of linear modular equations can be computed in a similar manner as well (using definition 3).

As shown by experimental results in Section 7, the black-box use of SMT solver Yices to obtain proofs of unsatisfiability is not efficient (as compared to the use of HNF). The main reason for this seems to be the structure of P. Even though the encoding used to obtain P is natural, it is difficult for algorithms used in Yices to decide P.